Twisted L-Functions and Monodromy

Nicholas M. Katz

Contents

Twisted L-Functions and Monodromy

Introduction

The present work grew out of an entirely unsuccessful attempt to answer some basic questions about elliptic curves over $\mathbb{Q}$. Start with an elliptic curve E over $\mathbb{Q}$, say given by a Weierstrass equation

$$E: \qquad y^2 = 4x^3 - ax - b,$$

with a, b integers and $a^3 - 27b^2 \neq 0$. By Mordell's theorem [Mor], the group $E(\mathbb{Q})$ of $\mathbb{Q}$-rational points is a finitely generated abelian group. The dimension of the $\mathbb{Q}$-vector space $E(\mathbb{Q})\otimes_{\mathbb{Z}}\mathbb{Q}$ is called the Mordell–Weil rank, or simply the rank, of E. Thus we get a function

$$\{(a,b) \text{ in } \mathbb{Z}^2 \text{ with } a^3 - 27b^2 \neq 0\} \to \{\text{nonnegative integers}\}$$

defined by

$$(a,b) \mapsto \text{the rank of the curve } y^2 = 4x^3 - ax - b.$$

It is remarkable how little we know about this function. For example, we do not know if this function is bounded, or if there exist elliptic curves over $\mathbb{Q}$ of arbitrarily high rank. For a long time, it seems to have been widely believed that this function was bounded. But over the past fifty years, cleverer and cleverer constructions, by Néron [Ner–10], Mestre [Mes–11, Mes–12, Mes–15], Nagao [Nag–20], Nagao–Kouya [Nag–Ko–21], Fermigier [Fer–22], and Martin–McMillen [Mar–McM–23 and Mar–McM–24], have given curves over $\mathbb{Q}$ with higher and higher rank. At this writing in September of 2000 the highest known rank is 24, and the present consensus is that there may well exist elliptic curves over $\mathbb{Q}$ of arbitrarily high rank.

We might then ask if at least we can say anything about the average rank of elliptic curves. What does this question mean? One naive but accessible formulation is this. Since $a^3 - 27b^2 \neq 0$, we might fix a nonzero integer Δ, and look first at the set Ell_Δ defined as

$$\mathrm{Ell}_\Delta := \{(a, b) \text{ in } \mathbb{Z}^2 \text{ with } a^3 - 27b^2 = \Delta\}.$$

Now for each nonzero Δ in $\mathbb{Z}$, the equation

$$X^3 - 27Y^2 = \Delta$$

itself is an elliptic curve over $\mathbb{Q}$. So it has only finitely many solutions (a, b) in integers, by a celebrated result of Siegel giving the finiteness of the number of integral points on an elliptic curve over $\mathbb{Q}$. So the set Ell_Δ is finite. For each integer $N > 0$ we take the union of the sets E_Δ for $0 < |\Delta| \le N$, and obtain the finite set

$$\mathrm{Ell}_{\le N} := \{(a,b) \text{ in } \mathbb{Z}^2 \text{ with } 0 < |a^3 - 27b^2| \le N\}$$

We now form the average

$$\mathrm{avrk}_{\le N} := (1/\#\mathrm{Ell}_{\le N})\Sigma_{(a,b) \text{ in } \mathrm{Ell}_{\le N}} (\text{rank of } y^2 = 4x^3 - ax - b),$$

which is a nonnegative real (in fact rational) number.

So now we have a sequence

$$N \to \mathrm{avrk}_{\leq N}$$

of nonnegative real numbers. We do not know if it has a limit. If it does, it would be reasonable to call its limit the average rank of elliptic curves over $\mathbb{Q}$. It is not even known (unconditionally, see [Bru] for conditional results on questions of this type) that the limsup of this sequence is finite.

For a long time, it was widely believed that the large N limit of $\mathrm{avrk}_{\leq N}$ does exist, and that its value is 1/2. Moreover, it was believed that each of the three auxiliary sequences of ratios

fraction of points in $\mathrm{Ell}_{\leq N}$ with rank 0,

fraction of points in $\mathrm{Ell}_{\leq N}$ with rank 1,

and

fraction of points in $\mathrm{Ell}_{\leq N}$ with rank ≥ 2,

has a limit, and that these limits are 1/2, 1/2, and 0 respectively.

Today it is still believed that each of these four sequences has a limit, but there is no longer agreement on what their limits should be. Some numerical experiments ([Brum–McG], [Fer–EE], [Kra–Zag], [Wa–Ta]) support the view that a positive percentage of elliptic curves have rank two or more, i.e., that the fourth limit is nonzero. On the other hand, the philosophy of Katz–Sarnak ([Ka–Sar, RMFEM, Introduction] and [Ka–Sar, Zeroes]) suggests that the limits are as formerly expected, and (hence) that the contradictory evidence is an artifact of too restricted a range of computation.

At this point, we must say something about the L–function L(s, E) of an elliptic curve over $\mathbb{Q}$, and about the Birch and Swinnerton–Dyer conjecture. The curve $E/\mathbb{Q}$ has "conductor" an integer $N = N_E \geq 1$ (whose exact definition need not concern us here) with the property that $E/\mathbb{Q}$ has "good reduction" at precisely the primes p not dividing N. For each such p we define an integer $a_p(E)$ by writing the number of $\mathbb{F}_p$–points on the reduction as $p + 1 - a_p(E)$. The L–function L(s, E) of $E/\mathbb{Q}$ is defined as an Euler product $\prod_p L_p(s,f)$, whose Euler factor $L_p(s, E)$ at each p not dividing N is

$$(1 - a_p(E)p^{-s} + p^{1-2s})^{-1}$$

(and with a recipe for the factors at the bad primes which need not concern us here). The Euler product converges absolutely for $\mathrm{Re}(s) > 2$, thanks to the Hasse estimate

$$|a_p(E)| \leq 2\mathrm{Sqrt}(p).$$

It is now known, thanks to work of Wiles [Wi], Taylor–Wiles [Tay–Wi], and Breuil–Conrad–Diamond–Taylor [Br–Con–Dia–Tay], that every elliptic curve $E/\mathbb{Q}$ is modular. What this means is that given $E/\mathbb{Q}$, with conductor $N = N_E$, there exists a unique weight two cusp form $f = f_E$ of weight two on the congruence subgroup $\Gamma_0(N)$ of $SL(2, \mathbb{Z})$ which is an eigenfunction of the Hecke operators T_p for primes p not dividing N, whose eigenvalues are the integers $a_p(E)$,

$$T_p f_E = a_p(E) f_E \text{ for every p not dividing N,}$$

whose q–expansion at the standard cusp $i\infty$ is q + higher terms, and which is not a modular form on $\Gamma_0(M)$ for any proper divisor M of N.

Now given **any** integer $N \geq 1$ and **any** weight two normalized newform f on $\Gamma_0(N)$, i.e., a cusp form f on $\Gamma_0(N)$ which is an eigenfunction of the Hecke operators T_p for primes p not dividing N, with eigenvalues denoted $a_p(f)$,

$$T_p f = a_p f,$$

whose q–expansion at i∞ is

$$\Sigma_{n\geq 1}\, a_n q^n,\ a_1 = 1,$$

and which is not a modular form on $\Gamma_0(M)$ for any proper divisor M of N, the L–function L(s, f) of f is defined to be the Mellin transform of f. Thus L(s, f) is the Dirichlet series

$$L(s, f) = \Sigma_{n\geq 1}\, a_n n^{-s}.$$

This Dirichlet series has an Euler product $\prod_p L_p(s,f)$ whose Euler factor $L_p(s, f)$ at each p not dividing N is

$$(1 - a_p p^{-s} + p^{1-2s})^{-1}.$$

The Euler product converges absolutely for Re(s) > 2. The function L(s, f) extends to an entire function, and when it is "completed" by a suitable Γ–factor, it satisfies a functional equation under $s \mapsto 2-s$. The precise result is this. One defines

$$\Lambda(s, f) := N^{s/2}(2\pi)^{-s}\Gamma(s)L(s, f).$$

Then Λ(s, f) is entire, and satisfies a functional equation

$$\Lambda(s, f) = \varepsilon(f)\Lambda(2-s, f),$$

where $\varepsilon(f) = \pm 1$ is called the sign in the functional equation.

It turns out that the Euler factors at the bad primes in L(s, E) are equal to those in $L(s, f_E)$, so we have the identity

$$L(s, E) = L(s, f_E).$$

This in turn shows that

$$\Lambda(s, E) := N^{s/2}(2\pi)^{-s}\Gamma(s)L(s, E)$$

extends to an entire function, and satisfies a functional equation

$$\Lambda(s, E) = \varepsilon(E)\Lambda(2-s, E),$$

with $\varepsilon(E)$ $(:= \varepsilon(f_E)) = \pm 1$.

The upshot of all this discussion is that L(s, E) is holomorphic at the point s=1, so it makes sense to speak of the order of vanishing of L(s, E) at the point s=1. The basic Birch and Swinnerton–Dyer conjecture for E/ℚ is the assertion that the rank of E/ℚ is the order of vanishing of L(s, E) at s=1. [We say "basic" because there is a refined version which interprets not only the order of vanishing as the rank, but also specifies the leading coefficient in the power series expansion of L(s, E) at s=1.] It is instructive to note that the conjecture was made thirty years before it was known in general that L(s, E) even made sense at s=1.

One calls the order of vanishing of L(s, E) at s=1 the "analytic rank" of E/ℚ, denoted $\mathrm{rank}_{an}(E)$:

$$\mathrm{rank}_{an}(E) := \text{order of vanishing of } L(s, E) \text{ at } s=1.$$

What we now know about the basic Birch and Swinnerton–Dyer conjecture can be stated all too briefly:

1) if L(1, E) is nonzero, then E has rank zero.

2) if L(s, E) has a simple zero at s=1, then E has rank one.

In other words, what we know is that

$$\mathrm{rank}_{an}(E) \le 1 \Rightarrow \mathrm{rank}(E) = \mathrm{rank}_{an}(E).$$

To emphasize how little we know, it is perhaps worth pointing out that we know neither the a priori inequality

$$\mathrm{rank}(E) \le \mathrm{rank}_{an}(E),$$

nor the opposite a priori inequality

$$\mathrm{rank}_{an}(E) \le \mathrm{rank}(E).$$

[In the "function field case", the analogue of the first a priori inequality holds trivially, cf. [Tate–BSD], [Shio].]

In all the numerical experiments concerning rank of which we are aware, it is the analytic rank rather than the rank which is calculated. Thus the relevance of these experiments to the rank of elliptic curves is conditional on the truth of the Birch and Swinnerton–Dyer conjecture.

A basic observation, due to Shimura (and related by him to Birch at the 1963 Boulder conference in the context of relating twists of modular forms and elliptic curves, cf. [Bir–St]), is that if the sign $\varepsilon(E)$ in the functional equation of L(s, E) is -1 [respectively $+1$], then L(s, E) has a zero of odd [respectively even] order at s=1. So we have the implication

$$\varepsilon(E) = -1 \Rightarrow \mathrm{rank}_{an}(E) \text{ is } \ge 1, \text{ and odd.}$$

If the Birch and Swinnerton–Dyer conjecture holds, then

$$\varepsilon(E) = -1 \Rightarrow \mathrm{rank}(E) \text{ is } \ge 1, \text{ and odd.}$$

On the other hand, if $\varepsilon(E)$ is $+1$, then rank(E) is forced to be even, so **if** the rank is nonzero, it is at least two. We should point out here that the parity consequence

$$\mathrm{rank}_{an}(E) \equiv \mathrm{rank}(E) \bmod 2$$

of the Birch and Swinnerton–Dyer conjecture remains a conjecture, sometimes called the Parity Conjecture [Gov–Maz].

The expectation that the average rank of elliptic curves over $\mathbb{Q}$ be 1/2 is based on three ideas: first, that the Birch and Swinnerton–Dyer conjecture holds for all $E/\mathbb{Q}$; second, that half the elliptic curves have sign $\varepsilon(E) = +1$, and half have sign $\varepsilon(E) = -1$; and third, that for most elliptic curves, the rank is the minimum, namely zero or one, imposed by the sign in the functional equation.

The recent conjecture of Katz–Sarnak [Ka–Sar, RMFEM, page 14] about the distribution of the low–lying zeroes of L(s, E) would, if true, make precise and quantify the third idea above, that for most elliptic curves, the rank is the minimum imposed by the sign of the functional equation. We refer to [Ka–Sar, RMFEM, 6.9 and 7.5.5] for the definitions and basic properties of the "eigenvalue location measures" $\nu(+,j)$ and $\nu(-,j)$, j = 1, 2, ... on $\mathbb{R}$. What is important for our immediate purpose is that these are all probability measures supported in $\mathbb{R}_{\geq 0}$ which are absolutely continuous with respect to Lebesgue measure.

In order to formulate the conjecture, we must assume the Riemann Hypothesis for the L–functions L(s, E) of all E/$\mathbb{Q}$, namely that all the nontrivial zeroes of L(s, E) (i.e., all the zeroes of Λ(s, E)) lie on Re(s) = 1. If L(s, E) has an even functional equation, its nontrivial zeroes occur in conjugate pairs $1 \pm i\gamma_{E,j}$ with $0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq \ldots$. If E has an odd functional equation, then s=1 is a zero of L(s, E), and the remaining nontrivial zeroes of L(s, E) occur in conjugate pairs $1 \pm i\gamma_{E,j}$ with $0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq \ldots$.

We then normalize the heights $\gamma_{E,j}$ of these zeroes according to the conductor N_E of E as follows. We define the normalized height $\tilde{\gamma}_{E,j}$ to be

$$\tilde{\gamma}_{E,j} := \gamma_{E,j}\log(N_E)/2\pi.$$

Now let us return to the set

$$\mathrm{Ell}_{\leq N} := \{(a,b) \text{ in } \mathbb{Z}^2 \text{ with } 0 < |a^3 - 27b^2| \leq N\}.$$

We then break up $\mathrm{Ell}_{\leq N}$ into two subsets

$$\mathrm{Ell}_{\leq N,\pm}$$

according to the sign in the functional equation of the L–function of the E/$\mathbb{Q}$ given by the corresponding Weierstrass equation. It is known to the experts, but nowhere in the literature, that both ratios

$$\#\mathrm{Ell}_{\leq N,\pm}/\#\mathrm{Ell}_{\leq N}$$

tend to 1/2 as $N \to \infty$.

Conjecture (compare [Ka–Sar, RMFEM, page 14]) The normalized heights of low–lying zeroes of L–functions of elliptic curves over $\mathbb{Q}$ are distributed according to the measures $\nu(\pm, j)$, in the following sense. For any integer $j \geq 1$, and for any compactly supported continuous $\mathbb{C}$–valued function h on $\mathbb{R}$, we can calculate the integrals $\int_{\mathbb{R}} h d\nu(\pm,j)$ as follows:

$$\int_{\mathbb{R}} h d\nu(-, j) = \lim_{N \to \infty} (1/\#\mathrm{Ell}_{\leq N,-}) \textstyle\sum_{E \text{ in } \mathrm{Ell}_{\leq N,-}} h(\tilde{\gamma}_{E,j}),$$

and

$$\int_{\mathbb{R}} h d\nu(+, j) = \lim_{N \to \infty} (1/\#\mathrm{Ell}_{\leq N,+}) \textstyle\sum_{E \text{ in } \mathrm{Ell}_{\leq N,+}} h(\tilde{\gamma}_{E,j}).$$

What is the relevance of this conjecture to rank? Take, for each real t > 0, a continuous function $h_t(x)$ on $\mathbb{R}$ which has values in the closed interval [0, 1], is supported in [–t, t], and takes the value 1 at the point x=0, for instance

$\_\_\_\wedge\_\_\_.$

By the absolute continuity of $\nu(\pm, j)$ with respect to Lebesgue measure, we have

$$|\int_{\mathbb{R}} h_t d\nu(\pm, j)| \to 0 \text{ as } t \to 0.$$

Choose N large enough that $\mathrm{Ell}_{\leq N,\varepsilon}$ is nonempty for both choices of sign ε. Denote by $\delta_0(x)$ the characteristic function of $\{0\}$ in $\mathbb{R}$. Notice that we have the trivial inequality $h_t(x) \geq \delta_0(x)$ for all real x. For the choice +, we have

$$(1/\#\mathrm{Ell}_{\leq N,+}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,+}} h(\tilde{\gamma}_{E,j})$$
$$\geq (1/\#\mathrm{Ell}_{\leq N,+}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,+}} \delta_0(\tilde{\gamma}_{E,j})$$
$$:= \text{fraction of E in } \mathrm{Ell}_{\leq N,+} \text{ with } \mathrm{rank}_{an}(E) \geq j.$$

For the choice −, the L−function automatically vanishes once at s=1, but that zero is not on our list $0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq ...$, so we have

$$(1/\#\mathrm{Ell}_{\leq N,-}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,-}} h(\tilde{\gamma}_{E,j})$$
$$\geq (1/\#\mathrm{Ell}_{\leq N,-}) \Sigma_{E \text{ in } \mathrm{Ell}_{\leq N,-}} \delta_0(\tilde{\gamma}_{E,j})$$
$$:= \text{fraction of E in } \mathrm{Ell}_{\leq N,-} \text{ with } \mathrm{rank}_{an}(E) \geq j+1.$$

Taking the limit as $N \to \infty$, and setting j = 1, we find

$$0 = \lim_{N \to \infty} \text{fraction of E in } \mathrm{Ell}_{\leq N,+} \text{ with } \mathrm{rank}_{an}(E) \geq 1,$$

and

$$0 = \lim_{N \to \infty} \text{fraction of E in } \mathrm{Ell}_{\leq N,-} \text{ with } \mathrm{rank}_{an}(E) \geq 2.$$

Therefore, if we assume in addition the Birch and Swinnerton−Dyer conjecture for all E/$\mathbb{Q}$, we find a precise sense in which a vanishingly small fraction of elliptic curves over $\mathbb{Q}$ have rank greater than that imposed by the sign in the functional equation.

As measures on $\mathbb{R}_{\geq 0}$, the $\nu(\pm, j)$ all have densities, and these densities are the restrictions to $\mathbb{R}_{\geq 0}$ of entire functions, cf. [Ka−Sar, RMFEM, 7.3.6, 7.5.5]. A significant difference between the two measures $\nu(-,1)$ and $\nu(+,1)$ is that the density of $\nu(-,1)$ vanishes to second order at the origin x=0, while that of $\nu(+,1)$ is $2 + O(x^2)$ near x=0, cf. [Ka−Sar, RMFEM, AG.0.3 and AG.0.5].

Thus the imposed zero of L(s, E) at s=1 for E of odd functional equation "quadratically repels" the next higher zero $1 + i\gamma_{E,1}$, while for E of even functional equation the point s=1 does not repel the next higher zero $1 + i\gamma_{E,1}$. This is presumably the phenomenon underlying the fact that in the numerical experiments cited above which call into question the "average rank = 1/2" hypothesis, what is found numerically is that about half the curves tested have odd sign, and essentially all of these have analytic rank one, while among the other half of the curves tested, among those with even sign, between twenty and forty percent have analytic rank two or more. What may be happening is that, because $\nu(-,1)$ quadratically repels the origin, while $\nu(+,1)$ does not repel the origin, in any given range of numerical computation, the data on ranks of curves of odd sign will look "better" than the data on ranks of curves of even sign ["better" in supporting the idea that elliptic curves over $\mathbb{Q}$ "try" to have as low a rank as their signs will allow].

An attractive and apparently "easier" question to study is this. Fix one elliptic curve $E/\mathbb{Q}$, with Weierstrass equation

$$E: \quad y^2 = 4x^3 - ax - b$$

and conductor N_E. For each squarefree integer D, one defines the quadratic twist E_D of E by D to be the elliptic curve over $\mathbb{Q}$ of equation

$$E_D: \quad Dy^2 = 4x^3 - ax - b,$$

or, equivalently (multiply the equation by D^3 and change variables to Dx, D^2y),

$$E_D: \quad y^2 = 4x^3 - aD^2x - bD^3.$$

Denote by χ_D the primitive quadratic Dirichlet character attached to the quadratic extension $\mathbb{Q}(\mathrm{Sqrt}(D))/\mathbb{Q}$. Thus for odd primes p not dividing D, we have

$$\chi_D(p) = 1 \text{ if D is a square in } \mathbb{F}_p, -1 \text{ if not.}$$

For all primes p which are prime to $2\times D\times N_E$, the a_p for E and for E_D are related by

$$a_p(E_D) = \chi_D(p)a_p(E).$$

The conductor of E_D divides (a power of 2)$\times D^2\times N_E$. If we take $D \equiv 1 \bmod 4$ and relatively prime to N, then the conductor of E_D is D^2N_E. For any D relatively prime to N, E_D has the sign in its functional equation related to that of E by the rule

$$\varepsilon(E_D) = \chi_D(-N_E)\varepsilon(E).$$

Denote by $f := f_E$ the weight two normalized newform attached to E. The normalized newform attached to E_D is $f\otimes\chi_D$, the unique weight two normalized newform of any level dividing a power of $2DN_E$ whose Hecke eigenvalues at primes not dividing $2DN_E$ are given by the rule $a_p(E_D) = \chi_D(p)a_p(E)$ above.

So having fixed $E/\mathbb{Q}$, we can now ask the same questions as above for the family of curves E_D. Thus for real $X > 0$, we look at the set

$$\mathrm{Sqfr}_{\le X} := \{\text{squarefree integers D with } |D| \le X\}.$$

On this set we have the function

$$D \mapsto \text{rank of } E_D.$$

We can ask whether as $X \to \infty$, the quantities

average of rank(E_D) over $\mathrm{Sqfr}_{\le X}$,

fraction of D in $\mathrm{Sqfr}_{\le X}$, with rank(E_D) = 0,

fraction of D in $\mathrm{Sqfr}_{\le X}$, with rank(E_D) = 1,

fraction of D in $\mathrm{Sqfr}_{\le X}$, with rank(E_D) $\ge$ 2,

have limits, and, if so, what they are. Or if not, what the limsup's might be. And a more refined version is to break $\mathrm{Sqfr}_{\le X}$ up according to the sign in the functional equation of $L(s, E_D)$ into two sets $\mathrm{Sqfr}_{\le X,\pm}$, and repeat the above questions over these sets. There are almost no unconditional results.

If we admit the truth of the Birch and Swinnerton–Dyer conjectures for all the twists E_D, then these are questions about the behavior at s=1 of the L–functions $L(s, f\otimes\chi_D)$ as D varies. Let us further assume the Riemann hypothesis for the L–functions L(s, f) attached to all weight two normalized newforms f on all $\Gamma_0(N)$. Then we can formulate the following conjecture.

Conjecture [Ka–Sar, Zeroes, II (b) and pg 21] Fix a weight two normalized newform f on any $\Gamma_0(N)$. Break up the set $\mathrm{Sqfr}_{\le X}$ according to the sign in the functional equation of $L(s, f\otimes\chi_D)$ into two subsets $\mathrm{Sqfr}_{\le X,\pm}$. [It is known that both the ratios

$$\#\mathrm{Sqfr}_{\le X,\pm} / \#\mathrm{Sqfr}_{\le X}$$

tend to 1/2 at $X \to \infty$.] Then the normalized heights $\tilde{\gamma}_{D,j}$ of the low–lying zeroes of the L–functions $L(s, f\otimes\chi_D)$ are distributed according to the measures $\nu(\pm, j)$, in the following sense. For any integer $j \ge 1$, and for any compactly supported continuous $\mathbb{C}$–valued function h on $\mathbb{R}$, we can calculate the integrals $\int_{\mathbb{R}} h d\nu(\pm,j)$ as follows.

$$\int_{\mathbb{R}} h d\nu(-, j) = \lim_{X \to \infty} (1/\#\mathrm{Sqfr}_{\le X,-}) \Sigma_{D \text{ in } \mathrm{Sqfr}_{\le X,-}} h(\tilde{\gamma}_{D,j}),$$

and

$$\int_{\mathbb{R}} h d\nu(+, j) = \lim_{X \to \infty} (1/\#\mathrm{Sqfr}_{\le X,+}) \Sigma_{D \text{ in } \mathrm{Sqfr}_{\le X,+}} h(\tilde{\gamma}_{D,j}).$$

Exactly as above, the truth of this conjecture for f_E gives us

$$0 = \lim_{X \to \infty} \text{fraction of D in } \mathrm{Sqfr}_{\le X,+} \text{ with } \mathrm{rank}_{an}(E_D) \ge 1,$$

and

$$0 = \lim_{X \to \infty} \text{fraction of D in } \mathrm{Sqfr}_{\le X,-} \text{ with } \mathrm{rank}_{an}(E_D) \ge 2.$$

So if we assume in addition the Birch and Swinnerton–Dyer conjecture for all the $E_D/\mathbb{Q}$, we find that as $X \to \infty$, 100 percent of the even twists have rank zero, that 100 percent of the odd twists have rank one, and that the average rank of all the twists is 1/2. That this should be so was first conjectured by Goldfeld [Go].

The numerical experiments so far seem to support this conclusion moderately well for odd twists, but poorly for even twists. Again, the fact that $\nu(-,1)$ quadratically repels the origin, while $\nu(+,1)$ does not repel the origin, may be "why" the numerical data so far is "better" for odd twists than for even twists.

We now turn to the situation for elliptic curves over function fields over finite fields. Thus let k be a finite field, C/k a proper smooth geometrically connected curve, K := k(C) its function field, and E/K an elliptic curve with nonconstant j–invariant. Then E/K "spreads out" to an elliptic curve over some dense open set U of C, say $\pi : \mathcal{E} \to U$. By the theory of the Néron model, if such a spreading out exists over a given open U, it is unique. Moreover, there is a largest such U, called the open set of good reduction for E/K. [Because E/K has nonconstant j–invariant, it does not have good reduction everywhere on C.] The finite set of closed points of C at which E/K has bad reduction will be denoted Sing(E/K). By the Néron–Ogg–Shafarevich criterion, the open set of good reduction can be described as follows. Pick a prime number ℓ invertible in K, pick some spreading out

$$\pi : \mathcal{E} \to U$$

of E/K, and form the lisse rank two sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell$ on U, which by Hasse [Ha] is pure of weight one. Denoting by $j : U \to C$ the inclusion, form the "middle extension" (:= direct image) sheaf $\mathcal{F} := j_*R^1\pi_*\overline{\mathbb{Q}}_\ell$ on C. This sheaf $\mathcal{F}$ on C is independent of the auxiliary choice of spreading out used to define it, and the open set of good reduction for E/K is precisely the largest open set on which $\mathcal{F}$ is lisse. Thus Sing(E/K) as defined above is equal to Sing($\mathcal{F}$), the set of points of C at which $\mathcal{F}$ is not lisse.

The L–function L(T, E/K) is defined to be the L–function of C with coefficients in $\mathcal{F}$, itself defined as the Euler product

$$L(T, \mathcal{F}) := \prod_x \det(1 - T^{\deg(x)}\mathrm{Frob}_x \mid \mathcal{F}_x)^{-1}$$

over the closed points x of C. At each point x of good reduction, the reduction of E/K at x is an elliptic curve $\mathbb{E}_x$ over the residue field $\mathbb{F}_x$, and

$$\det(1 - T\mathrm{Frob}_x \mid \mathcal{F}_x) = 1 - a_xT + (\#\mathbb{F}_x)T^2 \text{ in } \mathbb{Z}[T],$$

where a_x is the integer defined by the equation

$$a_x := 1 + \#\mathbb{F}_x - \#\mathbb{E}_x(\mathbb{F}_x).$$

Thus the local factors at the points of good reduction are visibly $\mathbb{Z}$–polynomials, independent of the auxiliary choice of ℓ. This is true also of the factors at the points of bad reduction [De–Constants, 9.8].

The cohomological expression for this L–function

$$L(T, \mathcal{F}) = \prod_{i=0,1,2}(\det(1 - T\mathrm{Frob}_k \mid H^i(C\otimes_k\overline{k}, \mathcal{F})))^{(-1)^{i+1}}$$

simplifies. Because E/K has nonconstant j–invariant, the middle extension sheaf $\mathcal{F}$ is geometrically irreducible when restricted to any dense open set of $C\otimes_k\overline{k}$ on which it is lisse [De–Weil II, 3.5.5]. This in turn implies that the groups H^i vanish for $i\neq 1$. Thus we end up with the identity

$$L(T, E/K) = L(T, \mathcal{F}) = \det(1 - T\mathrm{Frob}_k \mid H^1(C\otimes_k\overline{k}, \mathcal{F})).$$

By Deligne [De–WeII, 3.2.3], $H^1(C\otimes_k\overline{k}, \mathcal{F})$ is pure of weight two. Thus L(T, E/K) = L(T, $\mathcal{F}$) lies in $1 + T\mathbb{Z}[T]$ and has all its complex zeroes on the circle $|T| = 1/q$ (i.e., $L(q^{-s}, E/K)$ has all its zeroes on the line Re(s) = 1).

By the Mordell–Weil theorem, the group E(K) is finitely generated. The (basic) Birch and

Swinnerton–Dyer conjecture for E/K asserts that the rank of E(K), denoted rank(E/K), is the order of vanishing of L(T, E/K) at the point T = 1/q, q := #k, or, equivalently, that rank(E/K) is the multiplicity of 1 as generalized eigenvalue of Frob_k on the Tate–twisted group $H^1(C\otimes_k\bar{k}, \mathcal{F})(1)$. We call this multiplicity the analytic rank of E/K:

$$\mathrm{rank}_{an}(E/K) := \mathrm{ord}_{T=1}\det(1-T\mathrm{Frob}_k \mid H^1(C\otimes_k\bar{k}, \mathcal{F})(1)).$$

The group $H^1(C\otimes_k\bar{k}, \mathcal{F})(1)$ has a natural orthogonal autoduality <,> which is preserved by Frob_k, i.e., Frob_k lies in the orthogonal group $O := \mathrm{Aut}(H^1(C\otimes_k\bar{k}, \mathcal{F})(1), <,>)$. Now for any element A of any orthogonal group O, its reversed characteristic polynomial

$$P(T) := \det(1-AT)$$

satisfies the functional equation

$$T^{\deg(P)}P(1/T) = \det(-A)P(T),$$

the sign in which is det(−A). Taking for A the action of Frob_k on $H^1(C\otimes_k\bar{k}, \mathcal{F})(1)$, we have the identity P(qT) = L(T, E/K), so we find the functional equation of the L–function of E/K:

$$(qT)^{\deg(L)}L(1/q^2T, E/K) = \varepsilon(E/K)L(T, E/K),$$

where ε(E/K) is the sign

$$\varepsilon(E/K) = \det(-\mathrm{Frob}_k \mid H^1(C\otimes_k\bar{k}, \mathcal{F})(1)).$$

So just as in the number field case, we have the implications

$$\varepsilon(E/K) = -1 \Rightarrow \mathrm{rank}_{an}(E/K) \text{ is odd, and } \geq 1,$$

$$\varepsilon(E/K) = +1 \Rightarrow \mathrm{rank}_{an}(E/K) \text{ is even.}$$

In the function field case, we also have an a priori inequality

$$\mathrm{rank}(E/K) \leq \mathrm{rank}_{an}(E/K).$$

[But the "parity conjecture", the assertion that we have an a priori congruence

$$\mathrm{rank}(E/K) \equiv \mathrm{rank}_{an}(E/K) \bmod 2,$$

is not known in either the number field or the function field case.]

What about quadratic twists of a given E/K? To define these, we suppose that the field K has odd characteristic. Then E/K is defined by an equation

$$y^2 = x^3 + ax^2 + bx + c$$

where $x^3 + ax^2 + bx + c$ in K[x] is a cubic polynomial with three distinct roots in $\bar{K}$. For any element f in $K^\times$, the quadratic twist E_f/K is defined by the equation

$$fy^2 = x^3 + ax^2 + bx + c.$$

Pick any dense open set U in C over which E/K has good reduction, and over which the function f has neither zero nor pole. Then E_f/K also has good reduction over U, say $\pi_f : \mathcal{E}_f \to U$, and the lisse sheaf $R^1(\pi_f)_*\bar{\mathbb{Q}}_\ell$ on U is obtained from $R^1\pi_*\bar{\mathbb{Q}}_\ell$ by twisting by the lisse rank one Kummer sheaf $\mathcal{L}_{\chi_2(f)}$ on U:

$$R^1(\pi_f)_*\overline{\mathbb{Q}}_\ell = \mathcal{L}_{\chi_2(f)} \otimes R^1\pi_*\overline{\mathbb{Q}}_\ell$$

[Recall that χ_2 is the unique character of order two of $k^\times$, and $\mathcal{L}_{\chi_2(f)}$ is the character of $\pi_1(U)$ whose value on the geometric Frobenius Frob_x attached to a closed point x of U with residue field $\mathbb{F}_x$ is $\chi_2(N_{\mathbb{F}_x/k}(f(x)))$. This twisting formula is the sheaf–theoretic incarnation of the relation

$$a_x(E_f/K) = \chi_2(N_{\mathbb{F}_x/k}(f(x)))a_x(E/K),$$

itself the function field analogue of the number field formula

$$a_p(E_D) = \chi_D(p)a_p(E).]$$

So if we denote by $j : U \to C$, the sheaf $\mathcal{F}_f := j_*R^1(\pi_f)_*\overline{\mathbb{Q}}_\ell$ on C attached to E_f/K is related to the sheaf $\mathcal{F} := j_*R^1\pi_*\overline{\mathbb{Q}}_\ell$ on C attached to E/K by the rule

$$\mathcal{F}_f = j_*(\mathcal{L}_{\chi_2(f)} \otimes j^*\mathcal{F}).$$

And the L–function of E_f/K is thus

$$L(T, E_f/K) = L(T, \mathcal{F}_f) = \det(1 - T\mathrm{Frob}_k \mid H^1(C\otimes_k\overline{k}, \mathcal{F}_f)).$$

Thus when we start with a single elliptic curve E/K, and pick a prime number ℓ invertible in K, we get a geometrically irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on C. To the extent that we wish to study the **L–functions** of twists E_f/K (rather than the twists themselves, or their actual ranks) the only input data we need to retain is the sheaf $\mathcal{F}$. Indeed, once we have $\mathcal{F}$, the sheaf $\mathcal{F}_f$ attached to a twist E_f/K is constructed out of $\mathcal{F}$ by the rule

$$\mathcal{F}_f = j_*(\mathcal{L}_{\chi_2(f)} \otimes j^*\mathcal{F}),$$

for $j : U \to C$ the inclusion of any dense open set on which f is invertible and on which $\mathcal{F}$ is lisse.

In the case of twists of an $E/\mathbb{Q}$, we twisted by squarefree integers D, and for growing real $X > 0$ we successively averaged over the finitely many such D with $|D| \le X$. What is the function field analogue?

When the function field K is a rational function field $k(\lambda)$ in one variable λ, every element $f(\lambda)$ of $K^\times$ can be written as $f = g(\lambda)^2h(\lambda)$, with $h(\lambda)$ a polynomial in λ of degree $d \ge 0$ which has all distinct roots in $\overline{k}$ (i.e., h is a squarefree polynomial). This expression is unique up to $(g, h) \mapsto (\alpha g, \alpha^{-2}h)$ for some α in $k^\times$.

So in this case, we might initially try to look at twists of a given E by **all** squarefree polynomials in λ of higher and higher degree d. We might hope that for a given degree d of twist polynomial h, the L–functions $L(T, E_h/K)$ form some sort of reasonable family of polynomials in T. But the degree of $L(T, E_h/K)$ depends on more than just the degree of the squarefree h. It is also sensitive to the zeroes and poles of h at points of Sing(E/K), the set where E/K has bad reduction. For this reason, it is better to abandon the crutch of polynomials and their degrees, and rather impose in advance the behavior of the twisting function f in $K^\times$ at all the points of Sing(E/K).

Since we are doing quadratic twisting, the local geometric behavior at a point x in C of the twist E_f/K sees $\mathrm{ord}_x(f)$ only through its parity. Let us fix an effective divisor D on C and look only at functions f on C whose divisor of poles is exactly D, and which have d := deg(D) distinct zeroes (over $\bar{k}$), none of which lies in Sing(E/K)∩(C−D). We denote by

$$\mathrm{Fct}(C, d, D, \mathrm{Sing}(E/K)\cap(C-D)) \subset L(D)$$

this set of functions. Then the interaction between f and Sing(E/K) can be read entirely from the divisor D, in fact, from the parity of $\mathrm{ord}_x(D)$ at each point x in Sing(E/K). In particular, if we want to force local twisting at a given point x in C, in particular at a point in Sing(E/K), we have only to take an effective D which contains the point x with odd multiplicity. This formulation has the advantage of working equally well over a base curve C of any genus, whereas the polynomial formulation was tied to having $\mathbb{P}^1$ as the base.

The upshot is that if we fix an effective divisor D on C, then as f varies in the space

$$\mathrm{Fct}(C, d, D, \mathrm{Sing}(E/K)\cap(C-D)),$$

all the L−functions $L(T, E_f/K)$ have a common degree. It turns out there is a sheaf−theoretic explanation for this uniformity. For any effective D whose degree d satisfies $d \geq 2g+1$, the space

$$\mathrm{Fct}(C, d, D, \mathrm{Sing}(E/K)\cap(C-D))$$

is, in a natural way, the set of k−points of a smooth, geometrically connected k−scheme

$$X := \mathcal{Fct}(C, d, D, \mathrm{Sing}(E/K)\cap(C-D))$$

of dimension d + 1 − g. And there is a lisse $\overline{\mathbb{Q}}_\ell$−sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

on the space X, whose stalk $\mathcal{G}_f$ at a k−valued point

$$f \text{ in } X(k) = \mathrm{Fct}(C, d, D, \mathrm{Sing}(E/K)\cap(C-D))$$

is the cohomology group $H^1(C\otimes_k\bar{k}, \mathcal{F}_f)$, and whose local characteristic polynomial $\det(1 - T\mathrm{Frob}_{k,f} | \mathcal{G}_f)$ is given by

$$\det(1 - T\mathrm{Frob}_{k,f} | \mathcal{G}_f) = \det(1 - T\mathrm{Frob}_{k,f} | H^1(C\otimes_k\bar{k}, \mathcal{F}_f)) = L(T, E_f/K).$$

Moreover, the Tate−twisted sheaf $\mathcal{G}(1)$ is pure of weight zero, and has an orthogonal autoduality, which induces on each individual cohomology group $H^1(C\otimes_k\bar{k}, \mathcal{F}_f)(1)$ the orthogonal autoduality responsible for the functional equation of $L(T, E_f/K)$. And for each finite extension k_n/k of given degree n, the stalks of $\mathcal{G}$ at the k_n−valued points $X(k_n)$ encode the L−functions of twists defined over k_n.

In this way, questions about the (distribution of the zeroes of the) L−functions $L(T, E_f/K)$, as f varies in the space

$$X(k) = \mathrm{Fct}(C, d, D, \mathrm{Sing}(E/K)\cap(C-D)),$$

become questions about the sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

on X. Thanks to Deligne's equidistribution theorem [Ka–Sar, RMFEM, 9.2.6], we can answer many of these questions in terms of the geometric monodromy group G_{geom} attached to the sheaf $\mathcal{G}$.

For example, **if** the group G_{geom} is the full orthogonal group, we automatically get the following results on average analytic rank.

1) The average analytic rank over k_n of twists defined by f's in $X(k_n)$ tends to 1/2 as $n \to \infty$. [And hence the average rank has a limsup $\leq 1/2$ as $n \to \infty$.]

2) For each choice of $\varepsilon = \pm 1$, the fraction $\#X(k_n)_{\text{sign }\varepsilon}/\#X(k_n)$ of twists with sign ε in the functional equation tends to 1/2 as $n \to \infty$.

3) In the set $\#X(k_n)_{\text{sign }+}$, the fraction of twists with $\text{rank}_{an} = 0$ tends to 1 as $n \to \infty$. [And hence in the set $\#X(k_n)_{\text{sign }+}$, the fraction of twists with rank = 0 tends to 1 as $n \to \infty$.]

4) In the set $\#X(k_n)_{\text{sign }-}$, the fraction of twists with $\text{rank}_{an} = 1$ tends to 1 as $n \to \infty$. [And hence in the set $\#X(k_n)_{\text{sign }-}$, the fraction of twists with rank ≤ 1 tends to 1 as $n \to \infty$.]

Suppose we take a sequence of effective divisors D_ν on C whose degrees d_ν are strictly increasing. Then we get a sequence of smooth k–schemes

$$X_\nu := \mathcal{F}ct(C, D_\nu, d, \text{Sing}(E/K)\cap(C-D_\nu))$$

and, on each X_ν, a lisse sheaf $\mathcal{G}_\nu$, say of rank N_ν. The ranks N_ν tend to ∞ with ν. Suppose that for every large enough ν, the group G_{geom} for the sheaf $\mathcal{G}_\nu$ on X_ν is the full orthogonal group $O(N_\nu)$. Then for each choice of sign $\varepsilon = \pm 1$, and each choice of integer $j \geq 1$, we can obtain the eigenvalue location measure $\nu(\varepsilon, j)$ as the following (weak $*$) double limit: the large ν limit of the large n limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_\nu(k_n)_{\text{sign }\varepsilon}$.

It was with these applications in mind that we set out to prove that, at least in characteristic $p \geq 5$, as soon as the effective divisor D on C has degree d sufficiently large, then G_{geom} for $\mathcal{G}$ is the full orthogonal group. Unfortunately, this assertion is not always true. What is true is that G_{geom} is either the full orthogonal group O or the special orthogonal group SO, provided only that E/K has nonconstant j–invariant, and that

$$d \geq 4g+4, \text{ and}$$
$$2g - 2 + d > \text{Max}(2\#\text{Sing}(E/K)(\overline{k}), 144).$$

[If p=3, this result remains valid provided that the sheaf $\mathcal{F}$ attached to E/K is everywhere tamely ramified, a condition which is automatic in higher characteristic.]

We prove that G_{geom} is O if E/K has multiplicative reduction (i.e., unipotent local monodromy) at some point of Sing(E/K) which is not contained in D.

But there are cases where G_{geom} is SO rather than O. If E/K does **not** have unipotent local monodromy at **any** point of Sing(E/K), and if every point of Sing(E/K) which occurs in D does so with even multiplicity, then $\mathcal{G}$ has even rank, say N, and an analysis of local constants, using [De–

Constants, 9.5] shows that G_{geom} lies in SO(N) (and hence is equal to SO(N), for d large). cf. Theorem 8.5.7.

An example of an E/K with nonconstant j but with no places of multiplicative reduction, is the twisted (by $\lambda(\lambda-1)$) Legendre curve

$$y^2 = \lambda(\lambda-1)x(x-1)(x-\lambda)$$

over $k(\lambda)$, $k := \mathbb{F}_p$, p any odd prime, which has bad reduction precisely at 0, 1, ∞, but at each of these points the monodromy is

$$(\text{quadratic character})\otimes(\text{unipotent}).$$

In this example, it turns out (cf. Corollary 8.6.7) that if the characteristic p is 1 mod 4, then all the L–functions over all k_n have **even** functional equations. But, if p is 3 mod 4, then the L–functions over even [respectively odd] degree extensions k_n have even [respectively odd] functional equations!

The Legendre curve itself,

$$y^2 = x(x-1)(x-\lambda)$$

over $k(\lambda)$, has unipotent local monodromy at both 0 and 1. And so if we twist by polynomials $f(\lambda)$ in $k[\lambda]$ of any fixed degree $d \geq 146$, which have all distinct roots in $\overline{k}$ and are invertible at both 0 and 1, the resulting sheaf $\mathcal{G}_d$ on $X_d := \mathcal{F}ct(\mathbb{P}^1, d\infty, d, \{0,1\})$ has $G_{geom} = O(N_d)$, with N_d equal to 2d if d is even, and to 2d–1 if d is odd.

Now the Legendre curve makes sense over $\mathbb{Z}[1/2][\lambda, 1/\lambda(\lambda-1)]$, and the space X_d makes sense over $\mathbb{Z}[1/2]$. For each fixed $d \geq 146$, it makes sense to vary the characteristic p, and ask average rank questions about twists of the Legendre curve over $\mathbb{F}_p(\lambda)$ by points in $X_d(\mathbb{F}_p)$ as $p \to \infty$. We get the same answers as we got by fixing p and looking at twists by points in $X_d(\mathbb{F}_p n)$ as $n \to \infty$. If we vary d as well, we can recover the eigenvalue location measures $\nu(\varepsilon,j)$ as well. For each choice of sign ε and integer $j \geq 1$, we can obtain the eigenvalue location measure $\nu(\varepsilon, j)$ as the following (weak $*$) double limit: the large d limit of the large p limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_d(\mathbb{F}_p)_{\text{sign } \varepsilon}$.

But there are some basic things we don't know, "even" about this Legendre example, and "even" in equal characteristic p. For example, it is easy to see that for any fixed p, $\#X_d(\mathbb{F}_p) \to \infty$ as $d \to \infty$. [Indeed, an element of $X_d(\mathbb{F}_p)$ is a degree d polynomial $f(\lambda)$ in $\mathbb{F}_p[\lambda]$ with all distinct roots in $\overline{\mathbb{F}}_p$, which is nonzero at the points 0 and 1. For $d \geq 3$, any **irreducible** polynomial of degree d in $\mathbb{F}_p[\lambda]$ will lie in $X_d(\mathbb{F}_p)$. And the number of degree d irreducibles in $\mathbb{F}_p[\lambda]$ is at least

$$(p-1)(1/d)(p^d - (d/2)p^{d/2}).]$$

It is also easy to see that for each choice of sign ε, the ratio

$$\#X_d(\mathbb{F}_p)_{\text{sign } \varepsilon}/\#X_d(\mathbb{F}_p)$$

tend to 1/2 as $d \to \infty$. [For d even, use [De–Const, 9.5] as in 8.5.7. For d odd, use the fact that for α in $\mathbb{F}_p^\times$ a nonsquare, and any f in $X_d(\mathbb{F}_p)$, the twists of the Legendre curve by f and by αf have opposite signs in their functional equations, cf. 5.5.2, case 3).] But for p fixed, we do **not know** any of the following 1) through 4).

1)The average rank of twists defined by f's in $X_d(\mathbb{F}_p)$ tends to 1/2 as $d \to \infty$.
2) In the set $X_d(\mathbb{F}_p)_{\text{sign} -}$ the fraction of twists with $\text{rank}_{an} = 1$ tends to 1 as $d \to \infty$.
3) In the set $X_d(\mathbb{F}_p)_{\text{sign} +}$ the fraction of twists with $\text{rank}_{an} = 0$ tends to 1 as $d \to \infty$.
4) For each choice of sign ε and integer $j \geq 1$, the eigenvalue location measure $\nu(\varepsilon, j)$ is the following (weak $*$) **single** limit: the large d limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_d(\mathbb{F}_p)_{\text{sign } \varepsilon}$.

Let us now stand back and see what ingredients were required in the above discussion of quadratic twists of E/K, an elliptic curve over a function field with a nonconstant j–invariant. The function field K is the function field of a projective, smooth, geometrically connected curve C/k, k a finite field. Over some dense open set U in C, E/K spreads out to an elliptic curve $\pi : \mathcal{E} \to U$. We fix a prime number ℓ invertible in k, and form the lisse sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell$ on U. It is lisse of rank two, pure of weight one, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-1)$. The assumption that the j–invariant is nonconstant is used only to insure that $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is geometrically irreducible on U. If k has characteristic $p \geq 5$, then $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is everywhere tamely ramified: this is the only way the hypothesis $p \geq 5$ is used. Denoting by $j : U \to C$ the inclusion, we form the sheaf

$$\mathcal{F} := j_*R^1\pi_*\overline{\mathbb{Q}}_\ell$$

on C. We then fix an effective divisor D on C of large degree. We form the quadratic twists E_f/K of E/K by variable f in L(D) which have deg(D) distinct zeroes (over $\overline{k}$), none of which lies in D or in $\mathrm{Sing}(\mathcal{F})\cap(C-D)$. The L–functions of these quadratic twists are the local L–functions of a lisse $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

at the k–points of a smooth, geometrically connected k–scheme

$$X := \mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

of dimension $d + 1 - g$.

The original elliptic curve E/K occurs **only** through the geometrically irreducible middle extension sheaf $\mathcal{F}$ on C. Once we have $\mathcal{F}$, we can forget where it came from! Our fundamental result in the elliptic case is the determination of the geometric and arithmetic monodromy groups attached to the lisse $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F})$$

on the smooth, geometrically connected k–scheme

$$X := \mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

of dimension $\deg(D) + 1 - g$.

In fact, we can study the L–functions of twists, by nontrivial tame characters χ of **any** order, of an **arbitrary** geometrically irreducible middle extension sheaf $\mathcal{F}$ on C. Again in this general setup, the L–functions of such twists are the local L–functions of a lisse $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$$

at the k–points of the same smooth, geometrically connected k–scheme

$$X := \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

of dimension deg(D) + 1 – g that occurred above for quadratic twists of elliptic curves. Again the question is to determine the arithmetic and geometric monodromy groups attached to $\mathcal{G}$.

The rank N of $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ grows with deg(D), indeed we have an a priori inequality

$$N := \mathrm{rank}\mathcal{G} \geq (2g - 2 + \deg(D))\mathrm{rank}(\mathcal{F}).$$

One case of our main technical result (Theorems 5.5.1 and 5.6.1) is this. Suppose that $\mathcal{F}$ is everywhere tamely ramified. Then for any effective divisor D of large degree, the geometric monodromy group G_{geom} for $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ is one of the following subgroups of GL(N):

O(N)
SO(N): possible only if N is even
Sp(N): possible only if N is even
a group containing SL(N).

We can be more precise about which cases arise for which input data $(\mathcal{F}, \chi)$. Unless χ has order two and $\mathcal{F}$ is self–dual on $C\otimes\overline{k}$, G_{geom} contains SL(N). If $\mathcal{F}$ is orthogonally self–dual on $C\otimes\overline{k}$, and χ has order two, then $\mathcal{G}$ is symplectically self–dual on $X\otimes\overline{k}$, and G_{geom} for $\mathcal{G}$ is Sp(N). If $\mathcal{F}$ is symplectically self–dual on $C\otimes\overline{k}$, and χ has order two, then $\mathcal{G}$ is orthogonally self–dual on $X\otimes\overline{k}$, and G_{geom} for $\mathcal{G}$ is either SO(N), possible only if N is even, or it is O(N).

We can drop the hypothesis that $\mathcal{F}$ be everywhere tame if we are in large characteristic (the exact condition is $p \geq \mathrm{rank}(\mathcal{F}) + 2$), and if we require in addition that the effective divisor D of large degree contain no point where $\mathcal{F}$ is wildly ramified. [This second condition is automatic for D's which are disjoint from the ramification of $\mathcal{F}$.]

Fix, then, input data $(\mathcal{F}, \chi, D)$ as above. As deg(D) grows, the sheaves $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ have larger and larger classical groups as their geometric monodromy groups. The general large N limit results of Katz–Sarnak [Ka–Sar, RMFEM] then give information about the statistical behaviour of the zeroes of the L–functions of the corresponding twists. This information always concerns a double limit $\lim_{\deg(D) \to \infty} \lim_{\deg(E/k) \to \infty}$. For each D we must consider, for larger and larger finite extensions E of k, the L–functions of all twists $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$ as f runs over the E–valued points X(E) of the parameter space

$$X = \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})\cap(C-D)).$$

We also work out some refinements of these results, where we change the inner limit. The first refinement is twist only by "primes" in X(E), i.e., by functions f in X(E) whose divisor of zeroes $\mathrm{div}_0(f)$ is a single closed point of $C\otimes_k E$. The terminology "prime" arises as follows. In the case when C is $\mathbb{P}^1$ and D is $d\infty$, an element f in X(E) is a polynomial f(t) in E[t] of degree d which has d distinct roots in $\overline{E}$ and which is invertible at the finite singularities of $\mathcal{F}$. Such an element f is "prime" if and only if f(t) is an irreducible polynomial in E[t]. More generally, we might twist only by f's in X(E) whose divisor of zeroes has any pre–imposed factorization pattern. For instance, we might twist only by f's in X(E) which "split completely" over E, i.e., by f's in X(E) which have d distinct zeroes in C(E).

A second refinement is to start not over a finite field, but over a ring of finite type over $\mathbb{Z}$, for instance over $\mathbb{Z}[1/N\ell]$. Then just as in the case of the Legendre family discussed above, we can look at twists by points in $X(\mathbb{F}_p)$ as $p \to \infty$. We get the same answers as we got by fixing p and looking at twists by points in $X(\mathbb{F}_p n)$ as $n \to \infty$. We can combine the two refinements. We can twist only by primes in $X(\mathbb{F}_p)$ as $p \to \infty$, or we can twist only by elements of $X(\mathbb{F}_p)$ which "split completely" over $\mathbb{F}_p$. Under mild hypotheses, the limit results remain the same.

Still working over $\mathbb{Z}[1/N]$, take a sequence of divisors D_ν whose degrees d_ν are strictly increasing. We get thus a sequence of parameter spaces

$$X_\nu := \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})\cap(C-D))$$

over $\mathbb{Z}[1/N]$. We can recover the eigenvalue location measure (whichever of $\nu(\varepsilon,j)$ or $\nu(j)$ is appropriate to the situation being considered) as the following (weak $*$) double limit: the large ν limit of the large p limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_\nu(\mathbb{F}_p)$.

If we fix the prime p, and let $\nu \to \infty$, then just as in the Legendre case discussed above, it is natural to ask if we can recover the eigenvalue location measure, whichever of $\nu(\varepsilon,j)$ or $\nu(j)$ is appropriate, as the following (weak $*$) single limit: the large ν limit of the distribution of the j'th normalized zero of the L–functions attached to variable points in $X_\nu(\mathbb{F}_p)$.

Let us now backtrack, and describe the logical organization of this book. It falls naturally into four parts:

Part I (Chapters 1, 2, 3, 4): background material, used in Part II.
Part II (Chapter 5) twisting, done over an algebraically closed field
Part III (Chapters 6, 7, 8): twisting, done over a finite field
Part IV (Chapters 9, 10): twisting, done over schemes of finite type over $\mathbb{Z}$.

The first chapter is devoted to results from representation theory. Its main result is Theorem 1.5.1, which depends essentially upon a beautiful result of Zarhin about recognizing when an irreducible Lie subalgebra of End(V), V a finite–dimensional $\mathbb{C}$–vector space, is either Lie(SL(V)) or Lie(SO(V)) or, if dim(V) is even, Lie(Sp(V)). It also requires a remarkable recent result [Wales] of Wales concerning finite primitive irreducible subgroups G of GL(V) containing elements γ of type

$$\gamma := \mathrm{Diag}(\zeta, \zeta, ..., \zeta, 1, 1, ..., 1),$$

with ζ a primitive n'th root of unity, $n \geq 3$, which occurs with multiplicity r, $1 \leq r < \dim(V)$. Wales result is that $\dim(V) \leq 4r$.

Wales's inequality was conjectured in an earlier version of this manuscript, written at a time when less was known. It was known, by Blichfeld's 60° theorem [Blich–FCG, paragraph 70, Theorem 8, page 96], that the case $n \geq 6$ could not arise: no finite irreducible primitive subgroup of GL(V) contains such an element. [Blichfeldt's 60° theorem is that in a finite irreducible primitive subgroup G of $GL(N, \mathbb{C})$, if an element g in G has an eigenvalue α such that every other eigenvalue of γ is within 60° of α (on either side, including the endpoints), then g is a scalar.] A little–known result of Zalesskii showed that in the case n=5 we have $\dim(V) = 2r$. For n = 3 or n=4, there were only results for $r \leq 2$. For r=1, we have Mitchell's theorem [Mit], that a finite irreducible primitive subgroup of GL(V) containing a pseudoreflection of order $n > 2$ exists only if $\dim(V) \leq 4$. For r=2, Huffman–Wales prove that if n=4 then $\dim(V) \leq 4$, and if n=3 then $\dim(V) \leq 8$, cf. [Huf–Wa, Theorems 2 and 3 respectively].

In an appendix to Chapter 1, we explained at length the result [AZ.1] of Zalesskii, and made some conjectures about what might be true in general. Wales then proved the most optimistic AZ.6.2 of these conjectures. Because his manuscript itself refers to some of the results in the appendix, we have left the appendix unchanged, except to add a note saying that Wales has now proven AZ.6.2. We have, however, simplified the original statement and proof of Theorem 1.5.1 by making use of Wales's inequality.

In the second chapter, we use the general theory of Lefschetz pencils over an algebraically closed field to develop some basic facts about the geometry of curves, which were surely well known in the nineteenth century.

The third chapter is concerned with induction of group representations, and with giving algebro–geometric criteria for induced representations to have various properties (e.g., to be autodual, to be irreducible).

The fourth chapter is a brief review of "middle convolution" and its effect on local monodromy as developed in [Ka–RLS]. This material depends in an essential way on Laumon's work on Fourier transform.

After all these preliminaries, we turn to our subject proper in Chapter 5, which is the technical core of the book. We work over an algebraically closed field, and compute monodromy groups of twist sheaves, using as essential ingredients results of all the previous chapters. [In the earlier version of this manuscript, written before Wales's result, we could not twist by characters χ of order 4 or 6 unless the input sheaf $\mathcal{F}$ had rank at most 2.]

In Chapter 6, we explain how to formulate over a general base scheme the setup we considered in Chapter 5.

In Chapter 7, we work over a finite field, and extract the diophantine consequences of the monodromy results of Chapter 5. The essential ingredient here is the work of Deligne in [De–Weil II], both his purity theorem and his equidistribution theorem.

In Chapter 8, we give applications to average analytic rank of twists of a given elliptic curve. This leads us into a long discussion of whether the monodromy group in question is O or

SO, and leads us to some very nice examples.

In Chapter 9, we begin to work systematically over a base which is a scheme of finite type over $\mathbb{Z}$, rather than "just" a finite field. We also introduce the notion of twisting by a "prime." We prove an equidistribution theorem for primes in divisor classes, which was presumably well known in the late 1920's and 1930's to people like Artin, Hasse, and Schmidt, but for which we do not know a reference. We then analyze when twisting only by primes changes nothing as far as equidistribution properties. This leads us to a simple but useful case of Goursat's Lemma.

In Chapter 10, we give "horizontal" versions (i.e., over $\mathbb{F}_p$ as $p \to \infty$) of all the results we found earlier over a finite field k (where we worked over larger and larger extension fields of the given k).

It is a pleasure to thank Cheewhye Chin for his help in preparing the index. I respectfully dedicate this book to the memory of my teacher Bernard Dwork, to whom I owe so very much.

Chapter 1: "Abstract" Theorems of Big Monodromy

1.0 Two generalizations of the notion of pseudoreflection

(1.0.1) It will be convenient to introduce two generalizations of the notion of pseudoreflection. Suppose we are given a finite–dimensional vector space V over a field K. We write GL(V) for $\mathrm{Aut}_K(V)$, so long as there is no ambiguity about the field K. Recall that an element A in GL(V) is called a pseudoreflection if its space of fixed points, Ker(A−1), has codimension one in V, or, equivalently, if the quotient spaceV/Ker(A−1) has dimension one.

(1.0.2) Given an integer $r \geq 0$, and an element A in GL(V), we say that A has drop r if Ker(A−1) has codimension r in V. In other words,

(1.0.2.1) $$\text{drop of } A := \dim(V/\mathrm{Ker}(A-1)).$$

(1.0.3) Thus the only element of drop zero is the identity, and the elements of drop one are precisely the pseudoreflections. For A not the identity, we think of the drop of A as a measure of how nearly A resembles a pseudoreflection: the lower its drop, the more A resembles a pseudoreflection.

(1.0.4) A further property that any pseudoreflection A automatically satisfies is that it acts as a scalar on the quotient space V/Ker(A−1), simply because that space is one–dimensional.

(1.0.5) We say that an element A in GL(V) is quadratic of drop r if it has drop r and if in addition either $r = 0$ or the action of A on the quotient space V/Ker(A−1) is scalar, in which case which we call this scalar the scale of A. The terminology "quadratic" goes back to Thompson [Th–QP], and refers to the fact that, if $\dim(V) > r \geq 1$, the minimal polynomial of an A which is quadratic of drop r is a quadratic polynomial, namely $(T-1)(T-\mathrm{scale}(A))$. Conversely, given A in GL(V) whose minimal polynomial is $(T-1)(T-\lambda)$ for some λ in $K^\times$, A is a quadratic of drop $r = \dim(V/\mathrm{Ker}(A-1))$ and scale λ.

(1.0.6) Given a group I (we have in mind an inertia group), a K–linear representation ρ of I on V, and an integer $r \geq 0$, we say that ρ has drop r if, denoting by $V^I \subset V$ the subspace of I–invariant vectors in V, $\dim(V/V^I) = r$. We say that ρ is quadratic of drop r if either r=0 or if the action of I on V/V^I is scalar, in which case we call the linear character by which I acts on V/V^I the scale of ρ.

(1.0.7) If the group I is cyclic, with generator γ, then the drop, say r, of the representation ρ is equal to the drop of the element $\rho(\gamma)$, and the representation ρ is quadratic of drop r if and only if the element $\rho(\gamma)$ is quadratic of drop r.

(1.0.8) What happens for a more general group? Obviously, if ρ has drop r (resp. is quadratic of drop r), then for every element γ in I, $\rho(\gamma)$ has drop $\leq r$ (resp. $\rho(\gamma)$ is quadratic of drop $\leq r$). However, the converse is false in general: one cannot infer the drop of a representation just from looking at the drops of elements. The simplest example is the subgroup of $GL(2, \mathbb{Z})$ consisting of all 2×2 integer matrices $((\pm 1,n), (0,1))$ with n in $\mathbb{Z}$, in its standard representation std or the direct sum of std and a trivial representation of any size. Each element acts as a pseudoreflection or as the identity (i.e., has drop ≤ 1), but the representation has drop two. If we take the direct sum of k copies of such a representation, each element has drop $\leq k$, but the representation has drop 2k.

Another simple example is the diagonal subgroup Γ of $SL(2n+1, \mathbb{Z})$ in its standard representation std or in the direct sum of std and a trivial representation of any size. Every element in Γ has drop $\le 2n$, but the representation has drop $2n+1$.

1.1 Basic Lemmas on elements of low drop

Drop Lemma 1.1.1 Let K be a field, $r \ge 0$ an integer, M/K a vector space of dimension $m > 4r^2$, and C in GL(M) an element of drop r. Suppose there exists a tensor factorization of M as $V\otimes_K W$ with dim(V) = a, dim(W) = b, $a \le b$, and elements A in GL(V), B in GL(W) such that $C = A\otimes B$. Then A is scalar. If r=0, B is also scalar. If $r \ge 1$, then a divides r, and (hence) $a \le r$.

proof It suffices to prove the assertion after an arbitrary extension of the ground field, so we may reduce to the case when K is algebraically closed. Write C in Jordan form as a direct sum of scalars times unipotent Jordan blocks, say

$$C = \oplus_i (\lambda_i\otimes \mathrm{Unip}(d_i) \text{ on } M_i), \dim(M_i) \text{ denoted } d_i.$$

In this direct sum decomposition, compute Ker(C−1):

$$\mathrm{Ker}(C-1 \text{ on } M) = \oplus_i (\mathrm{Ker}(\lambda_i\otimes \mathrm{Unip}(d_i) - 1) \text{ on } M_i).$$

The kernel of $\lambda\otimes\mathrm{Unip}(d) - 1$ is zero for $\lambda \ne 1$, and is one−dimensional for $\lambda=1$. So we find

$$[\Sigma_{i \text{ with } \lambda_i = 1} (d_i -1)] \ + \ [\Sigma_{i \text{ with } \lambda_i \ne 1} d_i] = \mathrm{codim}\,\mathrm{Ker}(C-1) = r.$$

Looking only at the second bracketed term, we see that the total number (counting multiplicity) of eigenvalues of C which are not 1 is at most r.

So any list of at least r+1 eigenvalues of C contains the number 1, and any list of at least 2r+1 eigenvalues of C contains 1 as its majority listing. Fix an eigenvalue α of A. As C is $A\otimes B$, $\alpha\beta_i$ is an eigenvalue of C for each eigenvalue β_i of B. Notice that $b \ge 2r+1$ [because $ab = m > 4r^2$, and $b \ge a$, so if $b \le 2r$ then $ab \le b^2 \le 4r^2$]. Thus among the $\{\alpha\beta_i\}_{i=1 \text{ to } b}$, the most prevalent value is 1. This means that α is the most prevalent of the $1/\beta_i$. So **every** eigenvalue of A is α. Replacing A by $(1/\alpha)A$, and B by αB, we reduce to the case that A is unipotent.

Once A is unipotent, we next show it is semisimple. If not, then A has as a direct summand a Jordan block Unip(t) of size $t \ge 2$. Write the Jordan normal form of B:

$$B = \oplus_i \beta_i\otimes\mathrm{Unip}(n_i),$$

with integers $n_i \ge 1$.

Then $A\otimes B$ has a direct summand

$$\oplus_i \beta_i\otimes\mathrm{Unip}(t)\otimes\mathrm{Unip}(n_i).$$

Now in a single summand $\beta_i\otimes\mathrm{Unip}(t)\otimes\mathrm{Unip}(n_i)$, what is the codimension of the space of invariants? If $\beta_i \ne 1$, the invariants vanish, so the codimension is tn_i.

If $\beta_i= 1$, we claim the invariants in $\mathrm{Unip}(t)\otimes\mathrm{Unip}(n_i)$ are of dimension $\mathrm{Min}(t, n_i)$.

Lemma 1.1.2 The invariants in Unip(d)⊗Unip(e) have dimension Min(d, e).

proof By symmetry, we may assume $d \geq e$. The dual of Unip(d), being unipotent and indecomposable of dimension d, is again isomorphic to Unip(d), so the invariants in Unip(d)⊗Unip(e) are the equivariant maps from Unip(d) to Unip(e). Think of Unip(d) as $K[T]/(T-1)^d$ with the action of T. The equivariant maps become the K[T]–homomorphisms from $K[T]/(T-1)^d$ to $K[T]/(T-1)^e$. As $d \geq e$, we have

$$\mathrm{Hom}_{K[T]-\mathrm{mod}}(K[T]/(T-1)^d, K[T]/(T-1)^e)$$

$$= \mathrm{Hom}_{K[T]/(T-1)^d-\mathrm{mod}}(K[T]/(T-1)^d, K[T]/(T-1)^e),$$

and by "evaluation at 1" this last Hom group is just

$$= K[T]/(T-1)^e,$$

which has dimension $e := \mathrm{Min}(d, e)$. QED for the lemma

1.1.3 Remark on Lemma 1.1.2 If our ground field K has characteristic zero, then we know the Jordan decomposition of Unip(d)⊗Unip(e). If $d \geq e$, we have

(1.1.3.1) $\quad \mathrm{Unip}(d)\otimes\mathrm{Unip}(e) \cong \oplus_{j=1 \text{ to } e} \mathrm{Unip}(d + e - 2j).$

Since a single Jordan block has a one–dimensional space of invariants, the truth of Lemma 1.1.2 in characteristic zero is immediate from this formula.

The above formula 1.1.3.1 for the Jordan decomposition in characteristic zero of Unip(d)⊗Unip(e), $d \geq e$, results from the well–known formula for the tensor product of two symmetric powers $\mathrm{Symm}^a(\mathrm{std})$ and $\mathrm{Symm}^b(\mathrm{std})$, $a \geq b$, of the standard representation std of the algebraic group SL(2) over any field K of characteristic zero, according to which

(1.1.3.2) $\quad \mathrm{Symm}^a(\mathrm{std})\otimes\mathrm{Symm}^b(\mathrm{std}) \cong \oplus_{j=0 \text{ to } b} \mathrm{Symm}^{a+b-2j}(\mathrm{std}).$

One takes $a := d-1$, $b := e-1$, and uses the fact that for each integer $n \geq 0$, the standard upper unipotent element $\{(1,1), (0,1)\}$ in $SL(2, \mathbb{Z})$ acts as Unip(n+1) in $\mathrm{Symm}^n(\mathrm{std})$.

(1.1.4) We now return to the proof of the Drop Lemma 1.1.1. Thanks to the above Lemma 1.1.2, the codimension of the invariants, already in the direct summand

$$\oplus_i \beta_i \otimes \mathrm{Unip}(t)\otimes\mathrm{Unip}(n_i)$$

of A⊗B, is

$$\Sigma_{i \text{ with } \beta_i = 1} tn_i + \Sigma_{i \text{ with } \beta_i \neq 1} [tn_i - \mathrm{Min}(t, n_i)]$$

$$\geq \Sigma_i [tn_i - \mathrm{Min}(t, n_i)]$$

$$\geq \Sigma_i [tn_i - n_i]$$

$$= \Sigma_i (t-1)n_i \geq \Sigma_i n_i = b \geq 2r+1 > r,$$

contradiction.

Thus A is scalar, so it is $\mathbb{1}_a$, the a×a identity. Then C = A⊗B is the direct sum of a copies of B. So C−1 is the direct sum of a copies of B−1, and hence

$$r = \text{codim of Ker}(C-1) = a\times\text{codim of Ker}(B-1).$$

If r=0, we infer that B = $\mathbb{1}_b$, the b×b identity. If r ≥ 1, we infer that a | r, as required. QED for the drop lemma

(1.1.5) We will also require the Lie algebra version of the drop lemma above.

Drop Lemma, Lie algebra version 1.1.6 Let K be a field of characteristic zero, r ≥ 0 an integer, M/K a vector space of dimension $m > 4r^2$, C in End(M) and λ in K such that $C - \lambda$ has rank r as endomorphism of M (i.e., $\mathrm{Ker}(C-\lambda)$ has codimension r). Suppose there exists a tensor factorization of M as $V\otimes_K W$ with dim(V) = a, dim(W) = b, a ≤ b, and elements A in End(V), B in End(W) such that C = A⊗1 + 1⊗B. Then A is scalar. If r=0, B is also scalar. If r ≥ 1, then a divides r, and (hence) a ≤ r.

proof Extend scalars from K to the fraction field K((T)) of the power series ring K[[T]] in one variable T. Then $\exp(T(C-\lambda))$ has drop r, and exp(TC) = exp(TA)⊗exp(TB). Write $\exp(T(C-\lambda))$ as $\exp(-\lambda T)\exp(CT)$. Thus we have

$$\exp(T(C-\lambda)) = \exp(-\lambda T)\exp(TA)\otimes\exp(TB) = \exp(TA)\otimes\exp(T(B-\lambda)).$$

Now apply the drop lemma to conclude that exp(TA) is scalar, that if r = 0, then also $\exp(T(B-\lambda))$ is scalar, and that if r ≥ 1, then a | r. Differentiating exp(TA) and setting T=0, we find that A is scalar. If r=0, we find similarly that $B - \lambda$, and hence B, is scalar. QED

1.2 Tensor products and tameness at ∞

Lemma 1.2.1 Fix an algebraically closed field k and a prime number ℓ which is invertible in k. Suppose given an irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a dense open set $U \subset \mathbb{A}^1$, which is tame at ∞. Suppose that there exist lisse $\overline{\mathbb{Q}}_\ell$–sheaves $\mathcal{G}$ and $\mathcal{H}$ on U such that $\mathcal{F} \cong \mathcal{G}\otimes\mathcal{H}$. Then there exists a (unique) lisse, rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}$ on $\mathbb{A}^1$ such that $\mathcal{G}\otimes\mathcal{L}^{-1}$ is tame at ∞.

proof If char(k) = 0, take $\mathcal{L}$ = the constant sheaf $\overline{\mathbb{Q}}_\ell$, which is the unique lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}$ on $\mathbb{A}^1$.

If char(k) = p > 0, denote by P(∞) ⊂ I(∞) the wild inertia group. Denote by $\mathcal{F}(\infty)$, $\mathcal{G}(\infty)$, $\mathcal{H}(\infty)$ the I(∞)–representations attached to these sheaves. Because $\mathcal{F}(\infty)$ is trivial on P(∞), $\mathcal{G}(\infty)\otimes\mathcal{H}(\infty)$ is trivial on P(∞), and hence $\mathcal{G}(\infty)$ and $\mathcal{H}(\infty)$ are each scalar representations, by inverse $\overline{\mathbb{Q}}_\ell^\times$–valued characters χ and χ^{-1} of P(∞). The character χ is continuous on P(∞) and invariant under I(∞)–conjugation, simply because χ is the restriction to P(∞) of the $\overline{\mathbb{Q}}_\ell$–valued continuous central function on I(∞)

$$\gamma \mapsto (1/\mathrm{rank}(\mathcal{G}))\mathrm{Trace}(\gamma|\mathcal{G}(\infty)).$$

If we pick a topological generator γ^{tame} of the tame quotient $\mathrm{I}(\infty)^{\mathrm{tame}} \cong \prod_{\ell \text{ not } p} \mathbb{Z}_\ell(1)$, we get an isomorphism of I(∞) with the semidirect product $\mathrm{P}(\infty) \ltimes \langle\gamma^{\mathrm{tame}}\rangle \cong \mathrm{P}(\infty) \ltimes \mathrm{I}(\infty)^{\mathrm{tame}}$. Since χ on P(∞) is invariant by I(∞)–conjugation, we can extend χ to a continuous character $\tilde{\chi}$ of I(∞) by decreeing that $\tilde{\chi}(\gamma^{\mathrm{tame}}) = 1$. By continuity, χ on P(∞) has finite p–power order (cf. [Ka–Sar, RMFEM, 9.0.7]) and hence $\tilde{\chi}$ has finite p–power order on I(∞). [So in fact $\tilde{\chi}$ is the unique extension of χ to a character of finite p–power order on I(∞), since the ratio of any two such extensions is a tame character of finite p–power order of I(∞).] By the theory of the "canonical extension" [Ka–LG, 1.4.2], $\tilde{\chi}$ extends uniquely to a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}$ of rank one on $\mathbb{A}^1$. By construction, $\mathcal{G}\otimes\mathcal{L}^{-1}$ is tame at ∞. To see that $\mathcal{L}$ is the unique lisse sheaf on $\mathbb{A}^1$ with this property, notice that the P(∞)–representation attached to any such $\mathcal{L}$ must be the character χ. Hence the ratio of any two such $\mathcal{L}$ is lisse on $\mathbb{A}^1$ and tame at ∞, so trivial. QED

(1.2.2) Here is a variant of the above lemma, where we work on a curve of higher genus.

Lemma 1.2.3 Fix an algebraically closed field k and a prime number ℓ which is invertible in k. Let C/k be a proper smooth connected curve. Fix a point ∞ in C(k). Fix integers r ≥ 1 and m ≥ 0. Suppose given an irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a dense open set U ⊂ C–{∞}, which is tame at ∞. Suppose that there exist lisse $\overline{\mathbb{Q}}_\ell$–sheaves $\mathcal{G}$ and $\mathcal{H}$ on U such that $\mathcal{F} \cong \mathcal{G}\otimes\mathcal{H}$. Then there exists a lisse, rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}$ on C–{∞} such that $\mathcal{G}\otimes\mathcal{L}^{-1}$ is tame at ∞. If char(k) p = > 0, we may choose $\mathcal{L}$ to have finite p–power order.

proof Exactly as in the previous argument, we take $\mathcal{L}$ the constant sheaf if we are in characteristic zero, otherwise we extend χ uniquely to a continuous character $\tilde{\chi}$ of I(∞) of finite p–power order. This time, we appeal to Harbater ([Harb–Mod], cf. also [Ka–LG, 2.1.4]) to show the existence of a lisse, rank one $\mathcal{L}$ on C–{∞} extending $\tilde{\chi}$ and still having the same finite p–power order. QED

Remark 1.2.4 One essential difference between Lemmas 1.2.1 and 1.2.3 is that in the general case 1.2.3, the $\mathcal{L}$ is no longer unique, even if we insist that $\mathcal{L}$ have finite p–power order, as now the ratio of any two such $\mathcal{L}$ is a p–power order character of π_1(C). So if C has nonzero p–rank h, then for every integer r such that the order of χ divides p^r, there are p^{rh} possible $\mathcal{L}$'s of order dividing p^r. Only if the p–rank of C is zero do we get unicity of an $\mathcal{L}$ of p–power order. And if we drop the requirement that $\mathcal{L}$ have finite p–power order, then $\mathcal{L}$ is indeterminate up to a character of $\pi_1(\mathrm{C}–\{\infty\})^{\mathrm{tame}}$. Already taking only characters with values in $1+\ell\mathbb{Z}_\ell$ gives a $(\mathbb{Z}_\ell)^{2g}$ of indeterminacy.

1.3 Tensor indecomposability of sheaves whose local monodromies have low drop

Theorem 1.3.1 Fix an algebraically closed field k and a prime number ℓ which is invertible in k. Suppose given an integer $r \geq 1$ and an irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a dense open set $U \subset \mathbb{A}^1$, which is tame at ∞. Suppose that at each finite singularity s of $\mathcal{F}$, I(s) acts with drop $\leq r$. Suppose that there exist lisse $\overline{\mathbb{Q}}_\ell$–sheaves $\mathcal{G}$ and $\mathcal{H}$ on U with $\mathrm{rank}(\mathcal{G}) \leq \mathrm{rank}(\mathcal{H})$ such that $\mathcal{F} \cong \mathcal{G}\otimes\mathcal{H}$. If $\mathrm{rank}(\mathcal{F}) > 4r^2$, then $\mathrm{rank}(\mathcal{G}) = 1$.

proof By Lemma 1.2.1, there exists a lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf on $\mathbb{A}^1$ such that $\mathcal{G}\otimes\mathcal{L}^{-1}$ is tame at ∞. So replacing $\mathcal{G}$ and $\mathcal{H}$ by $\mathcal{G}\otimes\mathcal{L}^{-1}$ and $\mathcal{H}\otimes\mathcal{L}$ respectively, we may assume in addition that $\mathcal{G}$ is tame at ∞.

Fix a geometric point u in U, and write $\pi_1(U)$ for $\pi_1(U, u)$. View $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ as continuous $\overline{\mathbb{Q}}_\ell$–representations of $\pi_1(U)$, denoted $\Lambda_{\mathcal{F}}$, $\Lambda_{\mathcal{G}}$, $\Lambda_{\mathcal{H}}$ respectively.

For any k–valued point s in $S := \mathbb{A}^1 - U$, and any element γ of $\pi_1(U)$ which lies in an inertia group I(s), we know that $\Lambda_{\mathcal{F}}(\gamma)$ has drop $\leq r$, and we have the tensor decomposition

$$\Lambda_{\mathcal{F}}(\gamma) = \Lambda_{\mathcal{G}}(\gamma)\otimes\Lambda_{\mathcal{H}}(\gamma).$$

Applying the Drop Lemma 1.1.1, we see that $\Lambda_{\mathcal{G}}$ is scalar on I(s), say with character ρ_s. By the theory of the "canonical extension" [Ka–LG, 1.5.6] applied with the points ∞ and 0 replaced by the points s and ∞, there exists a lisse, rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}_s$ on $\mathbb{A}^1 - \{s\}$ which is tame at ∞ and which at s gives the character ρ_s of I(s). So replacing $\mathcal{G}$ by $\mathcal{G}\otimes(\otimes_{s \text{ in } S}\, \mathcal{L}_s)^{-1}$, and replacing $\mathcal{H}$ by $\mathcal{H}\otimes(\otimes_{s \text{ in } S}\, \mathcal{L}_s)$, we may further reduce to the case where $\mathcal{G}$ is not only tame at ∞ but trivial on every finite inertia group I(s). Therefore $\mathcal{G}$ is trivial ($\mathbb{A}^1$ is tamely simply connected!). Then $\mathcal{F}$ is $\mathrm{rank}(\mathcal{G})$ copies of $\mathcal{H}$. As $\mathcal{F}$ is irreducible, $\mathrm{rank}(\mathcal{G})$ must be one. QED

(1.3.2) We now give a slight extension of the above result to the case of projective representations.

Theorem 1.3.3 Fix an algebraically closed field k and a prime number ℓ which is invertible in k. Suppose given an integer $r \geq 1$ and an irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a dense open set $U \subset \mathbb{A}^1$, which is tame at ∞. Suppose that at each finite singularity s of $\mathcal{F}$, I(s) acts with drop $\leq r$. Fix a geometric point u in U, and write $\pi_1(U)$ for $\pi_1(U, u)$. View $\mathcal{F}$ as a continuous $\overline{\mathbb{Q}}_\ell$–representation of $\pi_1(U)$, denoted $\Lambda_{\mathcal{F}}$. Suppose that $\Lambda_{\mathcal{F}}$ as a projective representation of $\pi_1(U)$ has a tensor factorization $\mathcal{A}\otimes\mathcal{B}$ with $\mathcal{A}$ and $\mathcal{B}$ continuous projective $\overline{\mathbb{Q}}_\ell$–representations of $\pi_1(U)$, with $\dim(\mathcal{A}) \leq \dim(\mathcal{B})$. If $\mathrm{rank}(\mathcal{F}) > 4r^2$, then $\dim(\mathcal{A}) =1$.

proof Because U is an open smooth connected curve over an algebraically closed field, we have

$$H^2(\pi_1(U), \mathbb{Z}/d\mathbb{Z}) = H^2(U, \mathbb{Z}/d\mathbb{Z}) = 0 \text{ for every integer } d \geq 1.$$

Hence there is no obstruction to lifting a projective representation ρ: $\pi_1(U) \to PGL(d, \overline{\mathbb{Q}}_\ell)$ to a linear representation $\tilde{\rho}$: $\pi_1(U) \to SL(d, \overline{\mathbb{Q}}_\ell)$. Lift $\mathcal{A}$ and $\mathcal{B}$ to linear representations to SL, and interpret the lifts as lisse sheaves $\mathcal{G}$ and $\mathcal{H}$ on U. Then $\mathcal{F}$ and $\mathcal{G}\otimes\mathcal{H}$ are projectively equivalent linear representations. Therefore for some lisse rank one sheaf $\mathcal{L}$ on U, we have $\mathcal{F} \cong \mathcal{L}\otimes\mathcal{G}\otimes\mathcal{H}$. Now apply the previous theorem 1.3.1 to conclude that $\mathcal{L}\otimes\mathcal{G}$, and hence $\mathcal{G}$, has rank one. QED

Cautionary Remark 1.3.4 Theorem 1.3.1 and, a fortiori, Theorem 1.3.3 are both **false** if we drop the hypothesis that $\mathcal{F}$ be tame at ∞. Here are some examples to show this.
(1.3.4.1) Choose an integer $r \geq 1$ and an integer $g > 2r$. Pick a prime number $p \geq 2r+4$, and an algebraically closed field k of characteristic p. We will work on the affine line, with parameter t, over the field k. Fix a prime number $\ell \neq p$. We will construct Lie–irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaves $\mathcal{G}$ and $\mathcal{H}$ of ranks r and 2g respectively on a dense open set U of $\mathbb{A}^1$ whose tensor product $\mathcal{F} := \mathcal{G}\otimes\mathcal{H}$ is Lie–irreducible of rank $2gr > 4r^2$, such that all the finite local monodromy groups I(t), t in $\mathbb{A}^1 - U$, act on $\mathcal{F}$ with drop $\leq r$. We first describe the sheaf $\mathcal{G}$. Fix a nontrivial additive character

$$\psi : \mathbb{F}_p \to (\overline{\mathbb{Q}}_\ell)^\times.$$

Denote by $\mathcal{L}_\psi$ the corresponding Artin–Schreier sheaf on $\mathbb{A}^1$. Take for $\mathcal{G}$ the Fourier transform $FT_\psi(\mathcal{L}_{\psi(t^{r+1})})$. Thus $\mathcal{G}$ is lisse of rank r on $\mathbb{A}^1$, and its G_{geom} is given [Ka–MG, Theorem 19, applied with n = r+1] by

SL(r), if r is odd,
Sp(r), if r is even.

We next describe $\mathcal{H}$. Choose a monic polynomial f(x) in k[x] of degree 2g with 2g distinct roots, and consider the one–parameter family C_t of hyperelliptic curves of genus g given by

$$C_t : y^2 = f(x)(x-t).$$

Over the open set U of $\mathbb{A}^1$ where f(t) is invertible, the (complete nonsingular models of the) C_t fit together to form a proper smooth curve

$$\pi : C \to U,$$

and we take for $\mathcal{H}$ the lisse $\overline{\mathbb{Q}}_\ell$–sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell$ on U. By [Ka–Sar, RMFEM, 10.1.12–15], $\mathcal{H}$ is everywhere tame, all its finite monodromies have drop ≤ 1, and its G_{geom} is Sp(2g). By Goursat's lemma [Ka–ESDE, 1.8.2], G_{geom} for $\mathcal{G}\oplus\mathcal{H}$ is the product group

SL(r) × Sp(2g), if r is odd,
Sp(r)× Sp(2g), if r is even.

Therefore G_{geom} for $\mathcal{G}\otimes\mathcal{H}$ is the group

$SL(r) \times Sp(2g)$ if r is odd,
$(Sp(r)\times Sp(2g))/\pm(1,1)$, if r is even,

in its Lie–irreducible representation $(std_r)\otimes(std_{2g})$. Because $\mathcal{G}$ is lisse of rank r on all of $\mathbb{A}^1$, and each finite local monodromy of $\mathcal{H}$ has drop ≤ 1, each finite local monodromy of $\mathcal{G}\otimes\mathcal{H}$ has drop $\leq$ r.

(1.3.4.2) We can make even more egregious examples, by taking **both** $\mathcal{G}$ and $\mathcal{H}$ to be lisse on $\mathbb{A}^1$. Choose integers $r \geq 1$ and $m > 4r$. Pick a prime $p \geq 2m+4$. With ℓ and ψ chosen as in 1.3.4.1, take $\mathcal{G}$ to be the Fourier transform $FT_\psi(\mathcal{L}_{\psi(t^{r+1})})$, and take $\mathcal{H}$ to be the Fourier transform $FT_\psi(\mathcal{L}_{\psi(t^{m+1})})$. By [Ka–MG, Theorem 19, applied with n = r+1 and n = m+1 respectively], $\mathcal{G}$ [resp. $\mathcal{H}$] is lisse on $\mathbb{A}^1$ of rank r [resp. rank m], and its G_{geom} is the group SL(r) if r is odd, Sp(r) if r is even [resp. the group SL(m) if m is odd, Sp(m) if m is even]. Again using Goursat's lemma as in 1.3.4.1 above, we see that $\mathcal{F} := \mathcal{G}\otimes\mathcal{H}$ is Lie–irreducible on $\mathbb{A}^1$, of rank $rm > 4r^2$, and all the finite local monodromies of $\mathcal{F}$ act trivially, so have drop $0 \leq r$.

1.4 Monodromy groups in the Lie–irreducible case

Theorem 1.4.1 Fix an algebraically closed field k and a prime number ℓ which is invertible in k. Let C/k be a proper smooth connected curve, s a point in C(k). Fix an integer r with $r \geq 1$. Suppose given a Lie–irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a dense open set $U \subset C-\{s\}$, corresponding to a continuous $\overline{\mathbb{Q}}_\ell$–representation $\Lambda_\mathcal{F}$ of $\pi_1(U, u)$ on $V := \mathcal{F}_u$. Suppose that the action of I(s) on $\mathcal{F}$ is quadratic of drop r, and its scale is a linear character χ of I(s), possibly trivial, which is **not** of order 2. Then we have:

1) If χ is trivial, then G_{geom} contains a unipotent element A which is a quadratic of drop r, and $Lie(G_{geom})^{der}$ contains a nilpotent element n which, as endomorphism of V, has rank r. Moreover, $n^2 = 0$ in End(V).

2) If χ is nontrivial, then $((G_{geom})^0)^{der}$ contains a semisimple element A such that for some scalar λ in $\overline{\mathbb{Q}}_\ell^\times$, λA is quadratic of drop r, and $Lie(G_{geom})^{der}$ contains a semisimple endomorphism f of V with precisely two distinct eigenvalues, λ_1 and λ_2, such that $f-\lambda_1$ as endomorphism of V has rank r.

proof As $Lie(G_{geom})$ acts irreducibly on V, it is reductive, and we have a direct sum decomposition

$$Lie(G_{geom}) = Lie(G_{geom})^{der} \oplus (scalars)\cap Lie(G_{geom}),$$

with $Lie(G_{geom})^{der}$ a semisimple Lie subalgebra of End(V) which acts irreducibly on V. We can also describe $\mathrm{Lie}(G_{geom})^{der}$ as the traceless matrices, i.e., as the intersection of $\mathrm{Lie}(G_{geom})$ with $\mathrm{Lie}(SL(V))$.

We first prove 1). If χ is trivial, then I(s) acts by unipotent elements. As unipotent elements in GL(V) have pro-ℓ order, the action of the wild inertia group P(s) is trivial, and the action of I(s) factors through its tame quotient $I(s)^{tame}$. So any topological generator of $I(s)^{tame}$ acts as an element, say A, which is unipotent and quadratic of drop r, and this A is the required element of G_{geom}. If we put n := Log(A), we get a nilpotent element n of $\mathrm{Lie}(G_{geom})$ which, as endomorphism of V, has rank r, and satisfies $n^2 = 0$. As n is nilpotent, it has trace zero, so lies in $\mathrm{Lie}(G_{geom})^{der}$.

We next prove 2). If χ is nontrivial, then we can diagonalize the action of I(s). As χ is not of order 2, some γ in I(s) acts as the diagonal matrix

$$B := \mathrm{Diag}(\alpha, \alpha, ..., \alpha, 1, 1, 1, ..., 1),$$

with some $\alpha \neq \pm 1$ repeated r times, and 1 repeated $\mathrm{rank}(\mathcal{F}) - r$ times. Denote by K the subgroup of GL(V) generated by B. Then K, acting by conjugation on End(V), normalizes $\mathrm{Lie}(G_{geom})$, and hence it normalizes $\mathrm{Lie}(G_{geom})^{der}$, the intersection of $\mathrm{Lie}(G_{geom})$ with $\mathrm{Lie}(SL(V))$. Thus K normalizes $\mathrm{Lie}(G_{geom})^{der}$, a semisimple Lie subalgebra of End(V) which acts irreducibly on V. We now apply Gabber's "torus trick" [Ka–ESDE, 1.0], whose statement we recall:

Theorem 1.4.2 (Gabber). Let $\mathcal{G}$ be a semisimple Lie subalgebra of End(V) which acts irreducibly on V. Suppose that a diagonal subgroup K of GL(V) normalizes $\mathcal{G}$. Let $\chi_1, ..., \chi_n$ be the n characters of K defined by the diagonal matrix coefficients; i.e., $k = \mathrm{Diag}(\chi_1(k), ..., \chi_1(k))$ for k in K. Consider the "torus" $\mathcal{T}$ in End(V) consisting of those diagonal matrices $\mathrm{Diag}(X_1, ..., X_n)$ whose entries satisfy the conditions

$$\sum X_i = 0$$
$$X_i - X_j = X_k - X_m \text{ whenever } \chi_i/\chi_j = \chi_k/\chi_m \text{ on K}.$$

Then $\mathcal{T}$ lies in $\mathcal{G}$.

Applying Gabber's "torus trick" to our situation, and remembering that $\alpha \neq \pm 1$, we find that $\mathrm{Lie}(G_{geom})^{der}$ contains the torus of all diagonal matrices of trace zero of the form

$$\mathrm{Diag}(X, X, ..., X, Y, Y, Y, ..., Y),$$

X repeated r times and Y repeated $\mathrm{rank}(\mathcal{F}) - r$ times. Thus if we define

$$d := \mathrm{rank}(\mathcal{F}) - r,$$

then $\mathrm{Lie}(G_{geom})^{der}$ contains the element

$$\mathrm{Diag}(d, d, ..., d, -r, -r, -r, ..., -r),$$

which is the required "f", and the group $((G_{geom})^0)^{der}$ contains the one-dimensional torus

$$\mathrm{Diag}(t^d, t^d, ..., t^d, t^{-r}, t^{-r}, t^{-r}, ..., t^{-r}).$$

A general element (e.g., take t not a root of unity of order dividing r+d) of this torus is the required A. QED for 1.4.1.

Theorem 1.4.3 Fix an algebraically closed field k and a prime number ℓ which is invertible in k. Suppose given an integer $r \geq 1$ and a Lie–irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a dense open set $U \subset \mathbb{A}^1$, which is tame at ∞. Fix a geometric point u in U, and view $\mathcal{F}$ as a linear representation $\Lambda_{\mathcal{F}}$ of $\pi_1(U) := \pi_1(U, u)$ on $V := \mathcal{F}_u$. Suppose that at each finite singularity s of $\mathcal{F}$, I(s) acts with drop $\leq r$. Suppose that for some t in $\mathbb{P}^1 - U$, the action of I(t) on $\mathcal{F}$ is quadratic of drop R with $1 \leq R \leq r$, and its scale is a linear character χ of I(t), possibly trivial, which is **not** of order 2. Then we have

1) If $\mathrm{rank}(\mathcal{F}) > 4r^2$, $\mathrm{Lie}(G_{geom})^{der}$ is a simple Lie algebra.

2) If $\mathrm{rank}(\mathcal{F}) > \mathrm{Max}(4r^2, 72R^2)$, then $\mathrm{Lie}(G_{geom})^{der}$ is either Lie(SO(V)) or Lie(SL(V)) or, if dim(V) is even, Lie(Sp(V)).

3) If $\mathrm{rank}(\mathcal{F}) > \mathrm{Max}(4r^2, 72R^2)$, and if the scale χ of the action of I(t) is a nontrivial character, not of order 2, then $\mathrm{Lie}(G_{geom})^{der}$ is Lie(SL(V)), i.e., G_{geom} contains SL(V).

4) If $\mathrm{rank}(\mathcal{F}) > \mathrm{Max}(4r^2, 72R^2)$, then either G_{geom} contains SL(V), or G_{geom} is SO(V) or O(V) or, if dim(V) is even, Sp(V).

5) If R = 1, and $\mathrm{rank}(\mathcal{F}) > 4r^2$, then either G_{geom} contains SL(V), or dim(V) is even and G_{geom} is Sp(V). If in addition the scale χ of the action of I(t) is a nontrivial character, not of order 2, then G_{geom} contains SL(V).

6) Suppose that at some point t in $\mathbb{P}^1 - U$, some element of I(t) acts on $\mathcal{F}$ as a **reflection**. If $\mathrm{rank}(\mathcal{F}) > 4r^2$, then either G_{geom} contains SL(V), or G_{geom} is O(V).

proof We first prove 1). Let us denote $\mathrm{Lie}(G_{geom})^{der}$ by $\mathcal{G}$. Thus $\mathcal{G}$ is a semisimple Lie subalgebra of End(V) which acts irreducibly on V. We argue by contradiction. Suppose $\mathcal{G}$ is not simple. Then $\mathcal{G}$ is a product of some number $n \geq 2$ of simple Lie algebras $\mathcal{G}_i$, i=1 to n, and the faithful irreducible representation V of $\mathcal{G}$ is the tensor product of faithful irreducible representations V_i of the simple factors $\mathcal{G}_i$. Take any partition of the indexing set {1, ..., n} into two disjoint nonempty subsets $\mathcal{A}$ and $\mathcal{B}$. Let us denote by $\mathcal{G}_{\mathcal{A}}$ (respectively $\mathcal{G}_{\mathcal{B}}$) the product of the simple factors $\mathcal{G}_i$ with i in $\mathcal{A}$, (respectively i in $\mathcal{B}$) and by $V_{\mathcal{A}}$ (respectively $V_{\mathcal{B}}$) the tensor product of the V_i with i in $\mathcal{A}$ (respectively i in $\mathcal{B}$). Then $\mathcal{G}$ is $\mathcal{G}_{\mathcal{A}} \times \mathcal{G}_{\mathcal{B}}$, and V is $V_{\mathcal{A}} \otimes V_{\mathcal{B}}$. Thus $\mathcal{G}_{\mathcal{A}}$ (respectively $\mathcal{G}_{\mathcal{B}}$) is a Lie subalgebra of $\mathrm{End}(V_{\mathcal{B}})$ (resp. of $\mathrm{End}(V_{\mathcal{B}})$). At the expense of interchanging $\mathcal{A}$ and $\mathcal{B}$, we may assume that $\dim(V_{\mathcal{A}}) \leq \dim(V_{\mathcal{B}})$.

By parts 1) and 2) of the above result 1.4.1, we know that $\mathcal{G}$ contains an element f such that for some scalar λ, $f-\lambda$ has rank R, with $1 \le R \le r$. Let us write f according to the decomposition of $\mathcal{G}$ as $\mathcal{G}_{\mathcal{A}}\times\mathcal{G}_{\mathcal{B}}$, say $f = (f_{\mathcal{A}}, f_{\mathcal{B}})$. Viewing f, $f_{\mathcal{A}}$, and $f_{\mathcal{B}}$ as endomorphisms of V, $V_{\mathcal{A}}$, and $V_{\mathcal{B}}$ respectively, we have

$$f = f_{\mathcal{A}}\otimes 1 + 1\otimes f_{\mathcal{B}}.$$

Applying the Lie algebra form 1.1.6 of the drop lemma, we conclude that $f_{\mathcal{A}}$ is scalar, and that $\dim(V_{\mathcal{A}}) \mid R$.

In particular, we have $\dim(V_{\mathcal{A}}) \le R \le r$. Since $\dim(V) > 4r^2$, we have $\dim(V_{\mathcal{B}}) > 4r > \dim(V_{\mathcal{A}})$. Therefore, in any grouping of the tensor factors V_i of V into two clumps, $V_{\mathcal{A}}\otimes V_{\mathcal{B}} = V$, exactly one term $V_{\mathcal{A}}$ has (small) dimension dividing R, and on this term $f_{\mathcal{A}}$ is scalar. The other term $V_{\mathcal{B}}$ has (large) dimension $> 4r$. In particular, exactly one of the factors has dimension dividing R, and one does not.

We now claim there is one and only one i, say i_0, for which V_i has dimension **not** dividing R.

We first show that there is at least one index i_0 such that V_{i_0} has dimension not dividing R. For if not, then the factorization $V_i\otimes(\otimes_{j\neq i}V_j)$ shows that f_i is scalar on V_i, for every i. Hence f is scalar, in which case for any scalar λ, $f-\lambda$ has rank either 0 or dim(V), never R. Contradiction.

Thus there exists an index i_0 with V_{i_0} of dimension not dividing R. Take the factorization of V as $V_{i_0} \otimes (\otimes_{j\neq i_0}V_j)$. It must be the second factor $\otimes_{j\neq i_0}V_j$ whose dimension $\otimes_{j\neq i_0}\dim(V_j)$ divides R, and so all the V_j with $j \neq i_0$ have dimension dividing R.

The group $\pi_1(U)$ acts by conjugation on $\mathcal{G} = \mathrm{Lie}(G_{geom})^{der}$, compatibly with its action on V. Think of each V_i as a representation of $\mathcal{G}$. The collection of representations $\{V_i\}_i$ is intrinsically attached to the data $(\mathcal{G}, V)$, and from V_i we recover $\mathcal{G}_i$ as the image of $\mathcal{G}$ in $\mathrm{End}(V_i)$. Among the $\{V_i\}_i$ we have distinguished a particular V_{i_0}, the unique one whose dimension does not divide R. Therefore $\pi_1(U)$ fixes the isomorphism class of V_{i_0}. Thus $\pi_1(U)$ also fixes the isomorphism class of the complementary factor $\otimes_{j\neq i_0}V_j$. Thus we get projective representations $\mathcal{A}$ and $\mathcal{B}$ of $\pi_1(U)$ on $\otimes_{j\neq i_0}V_j$ and on V_{i_0} respectively, and the tensor product $\mathcal{A}\otimes\mathcal{B}$ of these projective representations is the projective representation of $\pi_1(U)$ on V attached to the given linear representation $\Lambda_{\mathcal{F}}$. In this tensor factorization, $\mathcal{A}$ has small dimension dividing R, and $\mathcal{B}$ has large dimension $\ge 4r$. Because $\mathrm{rank}(\mathcal{F}) > 4r^2$, and $\mathcal{F}$ is tame at ∞, we may apply the above Theorem 1.3.3 to infer that $\dim(\mathcal{A})$ is one. This means that $\otimes_{j\neq i_0}V_j$, and hence each V_j with $i \neq i_0$, has dimension one. But V_j is a faithful representation of a simple Lie algebra, so it must have dimension at least two. This contradiction shows that $\mathcal{G}$ is in fact simple.

To prove 2) once we know that $\mathcal{G}$ is simple, we have only to invoke the following striking result of Zarhin.

Theorem 1.4.4 [Zar–SLA, Theorem 6, its proof, and proof of Lemma 4] Over an algebraically closed field k of characteristic zero, let V be a faithful irreducible representation of a simple Lie algebra $\mathcal{G}$. Let $R \geq 1$ be an integer. View $\mathcal{G}$ as a Lie subalgebra of End(V), and suppose that there exists a scalar λ in k and an element f in $\mathcal{G}$ such that, viewing f as an endomorphism of V, we have $\mathrm{rank}(f-\lambda) = R$. If $\dim(V) > 72R^2$, then $\mathcal{G}$ is the Lie algebra of either SO(V) or SL(V) or, if dim(V) is even, of Sp(V).

We now prove 3). If the scale χ of the action of I(t) is not the trivial character, or a character of order 2, the proof of Theorem 1.4.1 shows that $\mathrm{Lie}(G_{\mathrm{geom}})^{\mathrm{der}}$ contains the element

$$\mathrm{Diag}(d, d, ..., d, -R, -R, -R, ..., -R),$$

with d repeated R times, $-R$ repeated d times, and $d := \dim(V) - R$. The eigenvalues of this element are not stable under $x \mapsto -x$ (because $\dim(V) = d+R > 4r^2 \geq 4R^2 \geq 4R$, so $d > R$). But the eigenvalues of any element of either Lie(SO(V)) or, if dim(V) is even, Lie(Sp(V)) acting on V are stable under $x \mapsto -x$.

It remains to prove 4). By 3), $((G_{\mathrm{geom}})^0)^{\mathrm{der}}$ is either SL(V) or SO(V), or, if dim(V) is even, Sp(V). If $((G_{\mathrm{geom}})^0)^{\mathrm{der}}$ is SL(V), there is nothing to prove.

If $((G_{\mathrm{geom}})^0)^{\mathrm{der}}$ is SO(V), then G_{geom} lies in the normalizer of SO(V) in GL(V). This normalizer is the group of orthogonal similitudes $GO(V) := \mathbb{G}_m O(V)$, so we have the inclusions

$$SO(V) \subset G_{\mathrm{geom}} \subset \mathbb{G}_m O(V).$$

We must show that the image of $\pi_1(U)$ lies in O(V). For then we get $SO(V) \subset G_{\mathrm{geom}} \subset O(V)$. As the index of SO(V) in O(V) is two, G_{geom} will then be either SO(V) or O(V). The sheaf $\mathcal{F}$ is lisse on the open set $U \subset \mathbb{A}^1$, and tame at ∞. The quotient $\pi_1(U)^{\text{tame at } \infty}$ is topologically normally generated by all the inertia groups I(s) at all the finite singularities s in $\mathbb{A}^1 - U$ of $\mathcal{F}$ (because $\mathbb{A}^1$ over an algebraically closed field is tamely simply connected). So it suffices to see that each I(s), s in $\mathbb{A}^1 - U$, lands in O(V) under the representation $\Lambda_{\mathcal{F}}$. Take an element γ in such an I(s), and denote by A its image under $\Lambda_{\mathcal{F}}$. We know that A has drop $\leq r$, and we know that there exists a scalar λ in $\overline{\mathbb{Q}}_\ell^\times$ such that λA lies in O(V). All but at most r of the eigenvalues of A are equal to 1, and hence all but at most r of the eigenvalues of λA are equal to λ. But given an element of O(V), all but at most two of its eigenvalues can be grouped into $[(\dim(V)-1)/2]$ pairs of inverses $\{\alpha_i, \alpha_i^{-1}\}$. Since λA has at most r eigenvalues not λ, at most r of these inverse pairs $\{\alpha_i, \alpha_i^{-1}\}$ have either member not λ. As

$$[(\dim(V)-1)/2] \geq [(4r^2)/2] = 2r^2 > r,$$

at least one of these inverse pairs $\{\alpha_i, \alpha_i^{-1}\}$ must be $\{\lambda, \lambda\}$. Thus $\lambda = \lambda^{-1}$, so $\lambda = \pm 1$. But λA lies in $O(V)$, so $\pm A$ lies in $O(V)$, so A lies in $O(V)$.

If $\dim(V)$ is even and $((G_{geom})^0)^{der}$ is $Sp(V)$, then G_{geom} lies in the normalizer of $Sp(V)$ in $GL(V)$. This normalizer is the group of symplectic similitudes $GSp(V) := \mathbb{G}_m Sp(V)$, so we have the inclusions

$$Sp(V) \subset G_{geom} \subset \mathbb{G}_m Sp(V).$$

Exactly as in the orthogonal case, it suffices to show that each $I(s)$, s in $\mathbb{A}^1 - U$, lands in $Sp(V)$ under the representation $\Lambda_{\mathcal{F}}$. This is shown exactly as in the orthogonal case, now using the fact that the eigenvalues of any element of $Sp(V)$ fall into $\dim(V)/2$ pairs of inverses $\{\alpha_i, \alpha_i^{-1}\}$.

To prove 5), we argue as follows. We are given that $\mathcal{F}$ is Lie–irreducible, so $\mathrm{Lie}(G_{geom})^{der}$ is an irreducible semisimple Lie subalgebra of $\mathrm{End}(V)$. Since $R = 1$, $\mathrm{Lie}(G_{geom})$ and hence its intrinsic subalgebra $\mathrm{Lie}(G_{geom})^{der}$ is normalized by a pseudoreflection which is not a reflection. By a result of Gabber [Ka–ESDE, 1.5], $\mathrm{Lie}(G_{geom})^{der}$ is either $\mathrm{Lie}(SL(V))$ or, if $\dim(V)$ is even, $\mathrm{Lie}(Sp(V))$. Now repeat the arguments given above for 3) and 4), which used only the inequality $\mathrm{rank}(\mathcal{F}) > 2r^2$.

The proof of 6) is similar to that of 5). Now $\mathrm{Lie}(G_{geom})^{der}$ is normalized by a reflection, and Gabber's result [Ka–ESDE, 1.5] tells us that $\mathrm{Lie}(G_{geom})^{der}$ is either $\mathrm{Lie}(SL(V))$ or $\mathrm{Lie}(SO(V))$. Now repeat the arguments given above for 3) and 4) to conclude that either G_{geom} contains $SL(V)$, or G_{geom} is $SO(V)$ or $O(V)$. Since G_{geom} contains a reflection, G_{geom} is not $SO(V)$. QED

1.5 Statement of the main technical result

Theorem 1.5.1 Fix an algebraically closed field k and a prime number ℓ which is invertible in k. Fix integers $r \geq 1$ and $m \geq 0$. Suppose given an irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a dense open set $U \subset \mathbb{A}^1$, say $U = \mathbb{A}^1 - S$. For each point t in $S\cup\{\infty\}$ at which the action of $I(t)$ is nontrivial and quadratic of drop $\leq r$, and with scale a character not of order 2, denote by R_t the drop at t. Define R_{min} to be the minimum of these R_t's. Define R_{min} to be $+\infty$, if there are no such points t. Suppose that $\mathcal{F}$ satisfies the following hypotheses 1) through 6):

1) $\mathcal{F}$ is tame at ∞.
2) At every s in S, the action of the inertia group $I(s)$ on $\mathcal{F}$ is nontrivial and has drop $\leq r$.
3) We have the inequality $m < \#S$.
4) There is a subset $S_0 \subset S$ with $\#S_0 \leq m$, such that for s in $S - S_0$, the action of $I(s)$ on $\mathcal{F}$ is

nontrivial and quadratic of drop $\le$ r, and its scale is a linear character of I(s), possibly trivial, which is **not** of order 2. [In particular, R_{min} is finite.]

5) Either (r+1)! is invertible in k, or $\mathcal{F}$ is tame at all points of S_0.

6) We have the inequality rank($\mathcal{F}$) > Max(2mr, $4r^2$, $72R_{min}{}^2$).

Pick a geometric point u in U, and view $\mathcal{F}$ as a continuous $\overline{\mathbb{Q}}_\ell$–representation $\Lambda_{\mathcal{F}}$ of $\pi_1(U)$:= $\pi_1(U, u)$ on V := $\mathcal{F}_u$. Denote by G_{geom} the Zariski closure of the image of $\pi_1(U)$ in GL(V). Then either G_{geom} contains SL(V), or G_{geom} is SO(V) or O(V), or, if dim(V) is even, Sp(V). Moreover, if at any point t in $\mathbb{P}^1$ – U, the action of I(t) is nontrivial and quadratic of some drop < rank($\mathcal{F}$), with scale a **nontrivial** character **not** of order 2, then G_{geom} contains SL(V).

1.6 Proof of Theorem 1.5.1

(1.6.1) It suffices to show that $\mathcal{F}$ is Lie–irreducible. For then, using hypotheses 1) through 4) and 6), the conclusion, except for the "moreover", results from Theorem 1.4.3 above. We deduce the "moreover" as follows. Suppose that at a point t in $\mathbb{P}^1$ – U, the action of I(t) is nontrivial and quadratic of drop < rank($\mathcal{F}$), with scale a nontrivial character not of order 2. Because the scale is a nontrivial character, I(t) and all elements in it act semisimply. Pick an element γ in I(t) such that γ^2 acts nontrivially. Then the element $\Lambda_{\mathcal{F}}(\gamma)$ in G_{geom} has exactly two distinct eigenvalues, 1 and some $\lambda \neq \pm 1$. But in the group O(V) and, if dim(V) is even, in the group Sp(V), all but at most two of the eigenvalues of any element can be grouped into [(dim(V) –1)/2] pairs of inverses $\{\alpha, \alpha^{-1}\}$, and the remaining one (in the case of O(odd)) or two (in the case of O(even)) are ±1. Since $\lambda \neq \pm 1$, no leftover eigenvalue can be λ. But neither $\{\lambda, \lambda\}$ nor $\{1, \lambda\}$ is a pair of inverses. So the element $\Lambda_{\mathcal{F}}(\gamma)$ cannot lie in either O(V) or Sp(V). So by the paucity of choice for G_{geom}, G_{geom} must contain SL(V).

(1.6.2) To show that $\mathcal{F}$ is Lie–irreducible, we use the general fact [Ka–MG] that an irreducible lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on a smooth open connected curve U over an algebraically closed field k in which ℓ is invertible is either Lie–irreducible, or is induced from a finite etale connected covering of U of degree d ≥ 2, or $\mathcal{F}$ is a tensor product $\mathcal{G}\otimes\mathcal{H}$ with $\mathcal{G}$ Lie–irreducible and $\mathcal{H}$ with finite monodromy and rank d ≥ 2. So we must show that $\mathcal{F}$ is neither induced nor a tensor product of type

(1.6.2.1) (Lie–irreducible)⊗(finite monodromy and rank ≥ 2).

(1.6.3) We first show that $\mathcal{F}$ is not induced from a finite etale connected covering of U of degree d ≥ 2. Here is the precise result.

Proposition 1.6.4 Notations as in Theorem 1.5.1, suppose that hypotheses 1) through 5) hold. If rank($\mathcal{F}$) > 2mr, $\mathcal{F}$ is not induced from a finite etale connected covering of U of degree d ≥ 2.

proof We argue by contradiction. Suppose that $\pi : V \to U$ is a finite etale covering of degree $d \geq 2$, with V connected, and $\mathcal{G}$ is a lisse $\overline{\mathbb{Q}}_\ell$–sheaf on V such that $\mathcal{F} = \pi_*\mathcal{G}$. Let us denote by X the complete nonsingular model of V, and by

$$\overline{\pi} : X \to \mathbb{P}^1$$

the finite flat map which prolongs π. Let us fix a point t in $\mathbb{P}^1 - U$, and denote by $x_1, ..., x_n$ the points of X lying over t. As representation of I(t), $\mathcal{F}(t)$ is $(\pi_*\mathcal{G})(t)$, which is the direct sum

$$\mathcal{F}(t) = \oplus_i \mathrm{Ind}_{I(x_i)}{}^{I(t)}\, \mathcal{G}(x_i).$$

Denote by K the function field of $\mathbb{P}^1$ over k, and by L the function field of X over k. Denote by K_t and L_{x_i} their completions at the indicated points, and by

$$\pi(x_i) : \mathrm{Spec}(L_{x_i}) \to \mathrm{Spec}(K_t)$$

the map induced on (the spectra of) these completions. Geometrically, we have

$$\mathcal{F}(t) = \oplus_i \pi(x_i)_*\mathcal{G}(x_i).$$

Lemma 1.6.4.1 The direct image $\pi(x_i)_*\mathcal{G}(x_i)$ is tame at t if and only if $\pi(x_i)_*\overline{\mathbb{Q}}_\ell$ is tame at t and $\mathcal{G}$ is tame at x_i. More precisely, we have

$$\mathrm{Swan}_t(\pi(x_i)_*\mathcal{G}(x_i)) = \mathrm{Swan}_{x_i}(\mathcal{G}) + \mathrm{rank}(\mathcal{G})\mathrm{Swan}_t(\pi(x_i)_*\overline{\mathbb{Q}}_\ell).$$

proof We will use a global argument. First, pick a second point $u \neq t$ in $\mathbb{P}^1$. By the theory of the canonical extension [Ka–LG, 1.4.1, but with t and u playing the roles of ∞ and 0], we can find a connected finite etale cover $f : Z \to \mathbb{P}^1 - \{u, t\}$ with Z connected, which is tame over u, and which over the punctured formal neighborhood $\mathrm{Spec}(K_t)$ of t is isomorphic to

$$\pi(x_i) : \mathrm{Spec}(L_{x_i}) \to \mathrm{Spec}(K_t).$$

Denote by x_i (sic!) the unique point of the complete nonsingular model $\overline{Z}$ lying over t. Pick a point y in $\overline{Z}$ lying over u in $\mathbb{P}^1$. By [Ka–LG, 2.1.6], we can find a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{G}_i$ on $\overline{Z} - \{x_i, y\}$ which is tame at y and for which $\mathcal{G}_i(x_i) \cong \mathcal{G}(x_i)$ as $I(x_i)$–representation. Now consider the virtual lisse sheaf of rank zero on Z given by $\mathcal{G}_i - \mathrm{rank}(\mathcal{G})\overline{\mathbb{Q}}_\ell$. Upstairs, the Euler–Poincaré formula gives

$$\begin{aligned}\chi(Z, \mathcal{G}_i - \mathrm{rank}(\mathcal{G})\overline{\mathbb{Q}}_\ell) &= -\textstyle\sum_{w \text{ in } \overline{Z} - Z} \mathrm{Swan}_w(\mathcal{G}_i - \mathrm{rank}(\mathcal{G})\overline{\mathbb{Q}}_\ell) \\ &= -\textstyle\sum_{w \text{ in } \overline{Z} - Z} \mathrm{Swan}_w(\mathcal{G}_i) \\ &= -\mathrm{Swan}_{x_i}(\mathcal{G}_i) \\ &= -\mathrm{Swan}_{x_i}(\mathcal{G}).\end{aligned}$$

But downstairs we have

$$\chi(Z, \mathcal{G}_i - \mathrm{rank}(\mathcal{G})\overline{\mathbb{Q}}_\ell) = \chi(\mathbb{P}^1 - \{u, t\}, f_*\mathcal{G}_i - \mathrm{rank}(\mathcal{G})f_*\overline{\mathbb{Q}}_\ell)$$
$$= -\mathrm{Swan}_t(f_*\mathcal{G}_i - \mathrm{rank}(\mathcal{G})f_*\overline{\mathbb{Q}}_\ell)$$

(there is no Swan_u by the imposed tameness of f and of $\mathcal{G}_i$ over u)

$$= -\mathrm{Swan}_t(\pi(x_i)_*\mathcal{G}(x_i) - \mathrm{rank}(\mathcal{G})\pi(x_i)_*\overline{\mathbb{Q}}_\ell).$$

Thus we get

$$\mathrm{Swan}_t(\pi(x_i)_*\mathcal{G}(x_i)) = \mathrm{Swan}_{x_i}(\mathcal{G}) + \mathrm{rank}(\mathcal{G})\mathrm{Swan}_t(\pi(x_i)_*\overline{\mathbb{Q}}_\ell). \text{ QED}$$

(1.6.4.2) We first apply the above Lemma 1.6.4.1 to $t=\infty$. We know that $\mathcal{F}(\infty)$ is tame, so we get that each local map $\pi(x_i)$ is tame, i.e., π is tame over ∞.

(1.6.4.3) We next show that the map π is tame, i.e., that Z/U is an everywhere tame covering. If $\mathcal{F}$ were everywhere tame, then we would get the tameness of π from the lemma above. In particular, if k has characteristic zero, then $\mathcal{F}$ is everywhere tame, and so π is tame.

Now we return to the general situation

$$\mathcal{F}(t) = \oplus_i\, \pi(x_i)_*\mathcal{G}(x_i).$$

If we take I(t) invariants $H^0(\mathrm{Spec}(K_t), ...)$, we get

$$\mathcal{F}(t)^{I(t)} = \oplus_i\, \mathcal{G}(x_i)^{I(x_i)}.$$

Thus we have

$$\mathcal{F}(t)/\mathcal{F}(t)^{I(t)} \cong \oplus_i\, \pi(x_i)_*\mathcal{G}(x_i)/\mathcal{G}(x_i)^{I(x_i)}.$$

Denote by $e_{i,t}$ the degree of L_{x_i}/K_t, i.e., $e_{i,t} = \deg(\pi(x_i))$. Then $\pi(x_i)_*\mathcal{G}(x_i)$ has rank equal to $e_{i,t}\mathrm{rank}(\mathcal{G})$, and $\mathcal{G}(x_i)^{I(x_i)}$ has rank at most $\mathrm{rank}(\mathcal{G})$. Thus we get

$$\mathrm{rank}(\pi(x_i)_*\mathcal{G}(x_i)/\mathcal{G}(x_i)^{I(x_i)}) \geq (e_{i,t} - 1)\mathrm{rank}(\mathcal{G}),$$

so an inequality

$$\mathrm{rank}(\mathcal{F}(t)/\mathcal{F}(t)^{I(t)}) \geq \mathrm{rank}(\mathcal{G})\textstyle\sum_i(e_{i,t} - 1).$$

Suppose now we take for t a point s of S_0. Then I(s) acts with drop $\leq$ r, so we get an inequality

$$r \geq \mathrm{rank}(\mathcal{F}(s)/\mathcal{F}(s)^{I(s)}) \geq \mathrm{rank}(\mathcal{G})\textstyle\sum_i(e_{i,s} - 1).$$

Therefore for each individual $e_{i,s}$ we have the inequality

$$r \geq e_{i,s} - 1.$$

If k has finite characteristic p, but (r+1)! is invertible in k, then p > r+1. Since p > r+1, we get $p > e_{i,s}$. Therefore the extension L_{x_i}/K_s has degree < p, so is tame. Thus π is tame over each point s in S_0.

If (r+1)! is not invertible in k, then by hypothesis 5b), $\mathcal{F}$ is tame at each point in S_0, and hence π is tame over each point of S_0. It remains to see that π is tame over each point of $S - S_0$. This results from the following lemma.

Lemma 1.6.4.3.1 The map π is finite etale over each point s in S at which the action of I(s) is nontrivial and quadratic, with scale a character χ_s of I(s) not of order two.

proof At such a point s, consider the decomposition

$$\mathcal{F}(s)/\mathcal{F}(s)^{I(s)} \cong \oplus_i \pi(x_i)_* \mathcal{G}(x_i)/\mathcal{G}(x_i)^{I(x_i)}.$$

Thus the action of I(s) on each summand $\pi(x_i)_*\mathcal{G}(x_i)/\mathcal{G}(x_i)^{I(x_i)}$ is scalar, by the character χ_s.

So the semisimplification $(\pi(x_i)_*\mathcal{G}(x_i))^{ss}$ of $\pi(x_i)_*\mathcal{G}(x_i)$ is a sum of copies of χ_s and of the trivial character $\mathbb{1}$. But induction from a subgroup of finite index commutes with semisimplification, so we have

$$\pi(x_i)_*(\mathcal{G}(x_i)^{ss}) = \text{a sum of copies of } \chi_s \text{ and of } \mathbb{1}.$$

For any representation $\mathcal{H}(x_i)$ of $I(x_i)$, $\mathcal{H}(x_i)$ is a direct factor of $\pi(x_i)^*\pi(x_i)_*(\mathcal{H}(x_i))$. Apply this to $\mathcal{G}(x_i)^{ss}$: we find that

$$\mathcal{G}(x_i)^{ss} = \text{a sum of copies of } \pi(x_i)^*\chi_s \text{ and of } \mathbb{1}.$$

If $\mathbb{1}$ is a summand of $\mathcal{G}(x_i)^{ss}$, then $\pi(x_i)_*\mathbb{1}$ (being a summand of $\pi(x_i)_*(\mathcal{G}(x_i)^{ss})$) is a sum of copies of χ_s and of $\mathbb{1}$, say

$$\pi(x_i)_*\mathbb{1} = a\mathbb{1} + b\chi_s.$$

Similarly, if $\pi(x_i)^*\chi_s$ is a summand of $\mathcal{G}(x_i)^{ss}$, then $\pi(x_i)_*\pi(x_i)^*\chi_s = \chi_s \otimes \pi(x_i)_*\mathbb{1}$ is a sum of copies of χ_s and of $\mathbb{1}$, and hence $\pi(x_i)_*\mathbb{1}$ is a sum of copies of $\mathbb{1}$ and χ_s^{-1}, say

$$\pi(x_i)_*\mathbb{1} = a\mathbb{1} + b\chi_s^{-1}.$$

Suppose first χ_s is nontrivial. Since χ_s does not have order 2, both χ_s and χ_s^{-1} take values not in $\mathbb{Z}$. But $\pi(x_i)_*\mathbb{1}$ is a permutation representation, so its trace has values in $\mathbb{Z}$. Therefore b=0, and $\pi(x_i)_*\mathbb{1} = a\mathbb{1}$. But the I(s)–invariants in $\pi(x_i)_*\mathbb{1}$ are the $I(x_i)$–invariants in $\mathbb{1}$, so are one–dimensional, and hence a=1. Thus $\pi(x_i)$ has degree one, as required.

If χ_s is trivial, then $\pi(x_i)_*\mathbb{1} = (a+b)\mathbb{1}$, and we conclude as above that $\pi(x_i)$ has degree one. QED for Lemma 1.6.4.3.1

(1.6.4.4) Thus the connected covering Z/U is everywhere tame, and is finite etale of degree d over $\mathbb{A}^1 - S_0$. Let us denote by $M \le m$ the number of points of S_0 over which Z is ramified, and by

$$s_1, s_2, \ldots, s_M,$$

the points themselves. The monodromy group, say G, of $\pi_*\overline{\mathbb{Q}}_\ell$ is a transitive (because Z is connected) subgroup of the symmetric group S_d. Because the covering is tame, its monodromy

group is generated by one element γ_s for each of the points s in $\mathbb{A}^1$ at which the covering is ramified. The conjugacy class in S_d of the element γ_s is simply described in terms of the ramification indices $e_{i,s}$ over s, as the product of disjoint cycles whose lengths are the $e_{i,s}$.
(1.6.4.5) Now think of G as sitting in S_d. How many of the symbols {1, 2, ..., d} do we use when we write out, as a product of disjoint cycles, one of its M generators γ_s? Cycles of length one aren't written, so we use precisely

$$\Sigma_{\text{i such that } e_{i,s} \geq 2}\, e_{i,s}$$

symbols. We have the inequality

$$\Sigma_{\text{i such that } e_{i,s} \geq 2}\, e_{i,s} \leq \Sigma_i\, 2(e_{i,s} - 1).$$

So each generator γ_s requires at most $2\Sigma_i(e_{i,s} - 1)$ of the symbols to write it.
(1.6.4.6) At each of the $M \leq m$ points s in question, we return to the inequality

$$r \geq \mathrm{rank}(\mathcal{F}(s)/\mathcal{F}(s)^{I(s)}) \geq \mathrm{rank}(\mathcal{G})\Sigma_i(e_{i,s} - 1),$$

which we rewrite as

$$2\Sigma_i(e_{i,s} - 1) \leq 2r/\mathrm{rank}(\mathcal{G}).$$

Thus each γ_s requires at most $2r/\mathrm{rank}(\mathcal{G})$ symbols to write it. Since there are $M \leq m$ generators, at most

$$2Mr/\mathrm{rank}(\mathcal{G}) \leq 2mr/\mathrm{rank}(\mathcal{G})$$

symbols are used in writing all the generators. But the subgroup of S_d these elements generate acts transitively, so certainly all of the symbols must be used in writing the generators (any unused symbol is fixed by every generator, hence by the entire group, contradicting transitivity). So we get

$$d \leq \#(\text{symbols used in writing generators}) \leq 2mr/\mathrm{rank}(\mathcal{G}).$$

Crossmultiplying, we find

$$\mathrm{rank}(\mathcal{F}) = d\times\mathrm{rank}(\mathcal{G}) \leq 2mr,$$

and this contradicts the hypothesis that $\mathrm{rank}(\mathcal{F}) > 2mr$. This contradiction shows that $\mathcal{F}$ is not induced, and concludes the proof of Proposition 1.6.4. QED

(1.6.5) We next show that $\mathcal{F}$ is not a tensor product of type

(Lie−irreducible)⊗(finite monodromy and rank ≥ 2).

Here is the precise result.

Proposition 1.6.6 Notations as in Theorem 1.5.1 above, suppose that hypotheses 1) through 5) hold, and that $\mathrm{rank}(\mathcal{F}) > \mathrm{Max}(2mr, 4r^2)$.
1) If $\mathcal{F}$ is a tensor product of type

(Lie−irreducible)⊗(finite monodromy and rank ≥ 2),

then $\mathcal{F}$ has finite monodromy which is irreducible and primitive.
2) $\mathcal{F}$ does not have finite monodromy which is irreducible and primitive. Hence, by 1), $\mathcal{F}$ is not a

tensor product of type

(Lie−irreducible)⊗(finite monodromy and rank ≥ 2).

proof 1) If $\mathcal{F}$ is a tensor product $\mathcal{G}\otimes\mathcal{H}$, then by Theorem 1.3.1 above, the smaller dimensional factor has dimension one. Since the finite monodromy factor has rank ≥ 2, we have $\mathcal{F} = \mathcal{L}\otimes\mathcal{H}$, with $\mathcal{L}$ of rank one and $\mathcal{H}$ with finite monodromy. Denote by $\Lambda_{\mathcal{F}}$, $\Lambda_{\mathcal{L}}$, and $\Lambda_{\mathcal{H}}$ the corresponding representations. We claim that $\mathcal{L}$ itself has finite monodromy, i.e, that the character $\Lambda_{\mathcal{L}}$ is of finite order. To see this, we argue as follows. Fix a point s in $S = \mathbb{A}^1 - U$. For an element γ in I(s), we have

$$\Lambda_{\mathcal{F}}(\gamma) = \Lambda_{\mathcal{L}}(\gamma)\otimes\Lambda_{\mathcal{H}}(\gamma).$$

The eigenvalues of $\Lambda_{\mathcal{F}}(\gamma)$ are thus $\Lambda_{\mathcal{L}}(\gamma)\times$\{the eigenvalues of $\Lambda_{\mathcal{H}}(\gamma)$\}. Denote by D the order of the finite image group $\Lambda_{\mathcal{H}}(\pi_1(U))$. Then every eigenvalue of $\Lambda_{\mathcal{H}}(\gamma)$ is a D'th root of unity, and hence every eigenvalue of $\Lambda_{\mathcal{F}}(\gamma)$ is of the form $\Lambda_{\mathcal{L}}(\gamma)\times$(a D'th root of unity). But $\Lambda_{\mathcal{F}}(\gamma)$ has drop ≤ r, so most of its eigenvalues are 1. Thus $\Lambda_{\mathcal{L}}(\gamma)$ is a D'th root of unity. Therefore $\mathcal{L}^{\otimes D}$ is lisse of rank one on all of $\mathbb{A}^1$, and hence has finite p−power order. [To see this, recall that $\Lambda_{\mathcal{L}}$ takes values in $\mathcal{O}_\lambda^\times$, for $\mathcal{O}_\lambda$ the ring of integers in some finite extension E_λ of $\mathbb{Q}_\ell$. Because the subgroup of finite index $1 + \ell\mathcal{O}_\lambda$ of $\mathcal{O}_\lambda^\times$ is pro−ℓ, $\Lambda_{\mathcal{L}}(P(\infty))$ is a finite p−group, say of order q. Then $\mathcal{L}^{\otimes Dq}$ is lisse on $\mathbb{A}^1$ and tame at ∞, so trivial.] Thus $\mathcal{L}$ is a character of finite order. Hence $\mathcal{F}$ itself has finite monodromy. By the previous proposition 1.6.4, $\mathcal{F}$ is not induced. Therefore the image $\Lambda_{\mathcal{F}}(\pi_1(U))$ is a finite irreducible primitive (not induced) subgroup of GL(V), and this finite group is equal to G_{geom}.

To prove 2), we argue by contradiction. Suppose then that G_{geom} is a finite irreducible primitive subgroup of GL(V). By hypotheses 3) and 4) of 1.5.1, there is a point t in S such that the action of I(t) is nontrivial and quadratic with scale a character, possibly trivial, whose order is not 2. The scale character cannot be trivial, otherwise G_{geom} contains a nontrivial unipotent element, contradicting its finiteness. The scale character cannot have infinite order, otherwise G_{geom} contains an element

$$\mathrm{Diag}(\alpha, ..., \alpha, 1, ..., 1)$$

with α not a root of unity, again contradicting its finiteness. [We use here again the fact that the scale character takes values in some $\mathcal{O}_\lambda^\times$, in which the group of roots of unity is finite. So if the scale character is of infinite order, it takes a value of infinite order.] Thus the scale character is nontrivial and has finite order, which by assumption is ≥ 3. So G_{geom} contains an element $\mathrm{Diag}(\zeta, ..., \zeta, 1, ...,1)$ with ζ a primitive n'th root of unity for some n ≥ 3, occurring with multiplicity R ≤ r. By Wales [Wales], the "most optimistic conjecture" AZ.6.2 holds, i.e., we have the inequality

$$\dim(V) \le 4R \le 4r.$$

But $\dim(V) = \mathrm{rank}(\mathcal{F})$, so this inequality contradicts the assumption that $\mathrm{rank}(\mathcal{F}) > 4r^2$. This concludes the proof of Proposition 1.6.6, and, with it, the proof of Theorem 1.5.1. QED

1.7 A sharpening of Theorem 1.5.1 when $R_{min} = 1$ or when some local monodromy is a reflection

Theorem 1.7.1 Notations as in Theorem 1.5.1, suppose either that
a) $R_{min} = 1$,
or
b) at some point t in $S\cup\{\infty\}$, some element of I(t) acts on $\mathcal{F}$ as a reflection.
Suppose that hypotheses 1) through 5) hold. Suppose further that

$$\mathrm{rank}(\mathcal{F}) > \mathrm{Max}(2mr, 4r^2).$$

In case a), either G_{geom} contains SL(V), or dim(V) is even and G_{geom} is Sp(V). In case b), either G_{geom} contains SL(V), or G_{geom} is O(V). Moreover, if at any point t in $\mathbb{P}^1 - U$, an element of I(t) acts as a pseudoreflection which is not unipotent, then G_{geom} contains SL(V).

proof Exactly as in the proof of Theorem 1.5.1, we use 1.6.4 and 1.6.6 to show that $\mathcal{F}$ is Lie–irreducible. Then we apply Theorem 1.4.3, part 5) to cover case a), and Theorem 1.4.3, part 6) to cover case b). QED

Appendix to Chapter 1: A Result of Zalesskii

The main results of this appendix are Propositions AZ.1, AZ.2, and AZ.4, all due to Zalesskii [Zal, 11.2].

Proposition AZ.1 Over $\mathbb{C}$, suppose G is a finite irreducible primitive subgroup of GL(V) which contains a quadratic element

$$\gamma := \mathrm{Diag}(\zeta, \zeta, ..., \zeta, 1, 1, ..., 1)$$

of drop r, $1 \le r < \dim(V)$. Suppose that ζ is a primitive fifth root of unity. Then $\dim(V) = 2r$.

proof Enlarge the group by adding to it the finite group μ_5 of scalars, i.e., replace G by $\mu_5 G$. This larger finite group contains G, so it acts irreducibly and primitively on V, and it contains the element

$$\zeta^2\gamma = \mathrm{Diag}(\zeta^3, \zeta^3, ..., \zeta^3, \zeta^2, \zeta^2, ..., \zeta^2).$$

So our result follows from

Proposition AZ.2 Over $\mathbb{C}$, suppose G is a finite irreducible primitive subgroup of GL(V) which contains an element

$$A := \mathrm{Diag}(\alpha, \alpha, ..., \alpha, \beta, \beta, ..., \beta)$$

with exactly two distinct eigenvalues, α and β, which are inverse primitive fifth roots of unity. Denote by $n(\alpha)$ and $n(\beta)$ the multiplicities of α and β as eigenvalues of A. Then α and β occur with equal multiplicity: $n(\alpha) = n(\beta)$.

proof Let G_1 be the normal subgroup of G generated by all the G–conjugates of A. Then V as a representation of G_1 must be isotypical, because V is an irreducible and noninduced representation of G. So $V|G_1$ is the direct sum of $k_1 \ge 1$ copies of an irreducible representation V_1 of G_1. Looking at the actions of A on V and on V_1, we see that the original multiplicities $n(\alpha)$ and $n(\beta)$ are both divisible by the integer k_1, and that A acting on V_1 has the same two eigenvalues α and β, but with multiplicities $n_1(\alpha) = n(\alpha)/k_1$ and $n_1(\beta) = n(\beta)/k_1$. That V_1 is not induced, i.e., that G_1 is a primitive irreducible subgroup of $GL(V_1)$, results from the following elementary lemma, applied to G_1 and V_1.

Lemma AZ.3 Over $\mathbb{C}$, suppose given a finite–dimensional vector space V. Suppose G is an irreducible subgroup of GL(V) which is generated by finitely many elements γ_i, each of which has the following property (***):
(***)given any eigenvalue α of γ_i, and given any integer $k \ge 2$, there exists a k'th root of unity ζ such that $\alpha\zeta$ is not an eigenvalue of γ_i.

Then G is a primitive irreducible subgroup of GL(V), i.e., the representation is not induced.

proof For an irreducible representation V of any group G, being induced is the same as having a direct sum decomposition ("system of imprimitivity") of V as $\oplus_i V_i$ into two or more nonzero subspaces such that for any g in G and any index i, there exists an index j such that g maps V_i to V_j. Expressed this way, it is clear that if we view G as a quotient of some other group Γ, and view V as a representation of Γ, then V is induced as a G–representation if and only if it induced as a Γ–representation.

Denote by n the number of generators γ_i, pick n distinct points t_i in $\mathbb{A}^1(\mathbb{C})$, and view G as a quotient of $\pi_1(\mathbb{A}^1(\mathbb{C}) - \{t_1, ..., t_n\})$, with

$$\text{a small loop around } t_i \mapsto \gamma_i.$$

View the representation V of G as a rank N := dim(V) $\mathbb{C}$–local system $\mathcal{F}$ on

$$U := \mathbb{A}^1(\mathbb{C}) - \{t_1, ..., t_n\},$$

whose local monodromy around t_j is γ_j. If V is induced as a G–representation, then $\mathcal{F}$ is induced from a connected finite etale covering π: Z → U of degree d ≥ 2. Thus $\mathcal{F}$ is $\pi_*\mathcal{G}$ for a local system $\mathcal{G}$ on Z. As $\mathbb{A}^1(\mathbb{C})$ is simply connected, the covering Z/U must be ramified above at least one of the points t_i, say over t_1. Denote by $x_1, ..., x_m$ the points of $\overline{Z}$ lying over t_1, and by e_i the ramification index of x_i over t_1. At least one of them is ≥ 2, say e_1. Then a small disc centered at x_1 is mapped by π to a small disc centered at t_1 in suitable local coordinates by the e_i'th power mapping $[e_i]$. Then $\mathcal{F}(t_1)$ contains $[e_1]_*\mathcal{G}(x_1)$ as a direct summand. In terms of the eigenvalues ρ_i of local monodromy group of $\mathcal{G}(x_1)$, those of $[e_1]_*\mathcal{G}(x_1)$ are all the e_1'th roots of the ρ_i. In particular, among the eigenvalues of γ_1, which is local monodromy of $\mathcal{F}(t_1)$, are all the e_1'th roots of the nonzero complex number ρ_1. As all of the e_1'th roots of ρ_1 occur, any of them violates the property (***) that γ_1 was supposed to satisfy. This contradiction shows that $\mathcal{F}$ is not induced, or, equivalently, that the representation V of G is not induced. QED

We now return to proving Proposition AZ.2. Passing from (G, V) to (G_1, V_1) simply divides the multiplicities by the same factor k, and keeps the primitivity.

We continue this process. Denote by G_2 the subgroup of G_1 generated by all the G_1–conjugates of A. Since G_2 is normal in G_1, and V_1 is not induced, the restriction to G_2 of the representation V_1 is isotypical, say $V_1|G_2$ is the direct sum of $k_2 \geq 1$ copies of an irreducible representation V_2 of G_2. Looking at the action of A in both V_1 and V_2, we see that it has the same two eigenvalues α and β, and that their multiplicities $n_1(\alpha)$ and $n_1(\beta)$ in V_1 are k_2 times their multiplicities $n_2(\alpha)$ and $n_2(\beta)$ in V_2. The lemma AZ.3 above shows that V_2 is not induced. So we may continue in this fashion. Define G_{i+1} to be the subgroup of G_i generated by all the G_i–conjugates of A. Since G_{i+1} is normal in G_i, and V_i is not induced, the restriction to G_{i+1} of the representation V_i is isotypical, say $V_i|G_{i+1}$ is the direct sum of $k_{i+1} \geq 1$ copies of an irreducible representation V_{i+1} of G_{i+1}. Looking at the action of A in both V_i and V_{i+1}, we see that it has the same two eigenvalues α and β, and that their multiplicities $n_i(\alpha)$ and $n_i(\beta)$ in V_1 are k_{i+1} times their multiplicities $n_{i+1}(\alpha)$ and $n_{i+1}(\beta)$ in V_{i+1}. Since G is finite, this descending chain of subgroups must stabilize: at some point we will have $G_i = G_{i+1}$. At this point, G_i is generated by all the G_i conjugates of A. So we are reduced to proving the following Proposition.

Proposition AZ.4 Over $\mathbb{C}$, suppose α and β are inverse primitive fifth roots of unity, and $n(\alpha)$ and $n(\beta)$ are strictly positive integers. Suppose G is a finite irreducible primitive subgroup of GL(V) which is generated by all the G–conjugates of a single element A in G, which in GL(V) is GL(V)–conjugate to the element

$$\mathrm{Diag}(\alpha, \alpha, ..., \alpha, \beta, \beta, ..., \beta),$$

in which α (resp. β) occurs with multiplicity $n(\alpha)$ (resp. $n(\beta)$).
Then $n(\alpha) = n(\beta)$.

proof We can find a G–conjugate of A, say B, which does not commute with A. For if not, A lies in the center of G, and both of its eigenspaces are G–stable, contradicting irreducibility. Now denote by $H \subset G$ the subgroup generated by A and B, and decompose V as a representation of H. By Blichfeldt's "two eigenvalue argument" [Blich–FCG, paragraph 103], any irreducible H–submodule of V has dimension ≤ 2, cf. [Zal, 11.1]. [Blichfeldt's two eigenvalue result is that, over $\mathbb{C}$, if H is a finite subgroup of GL(V) generated by two elements, each of which has at most two distinct eigenvalues, then any irreducible H–submodule of V has dimension at most two.] So we have

$$V|H = (\oplus_i W_i) \oplus (\oplus_j \chi_j),$$

where the W_i are two–dimensional irreducible H–modules, and the χ_j are one–dimensional H–modules. Notice for later use that each χ_j has order 1 or 5, since H is generated by elements of order 5. There are some W_i in the decomposition of V|H, because V is a faithful representation of H, and H is not abelian.

Acting on any W_i, both A and B are conjugate in $GL(W_i)$ to $\mathrm{Diag}(\alpha, \beta)$, but do not commute in $GL(W_i)$. For if either A or B were scalar, or if A and B commuted in $GL(W_i)$, W_i would not be irreducible.

So in order to show that $n(\alpha) = n(\beta)$, it suffices to show that there are no χ_j in V|H. For then $V|H = \oplus_i W_i$, and A has eigenvalues $\{\alpha, \beta\}$ in each W_i. We now give Zalesskii's argument for the absence of any χ_j's.

By Lemma AZ.3 above, W_i is not induced as a representation of H. Let us denote by H(i) the image of H in $GL(W_i)$. In fact, H(i) lies in $SL(W_i)$, since each of A and B is conjugate in $GL(W_i)$ to $\mathrm{Diag}(\alpha, \beta)$. Thus H(i) is a finite irreducible primitive subgroup of $SL(W_i)$ generated by two elements of order 5, each with the same eigenvalues α and β. Consider the image $\overline{H}(i)$ in $PSL(W_i)$. It is not dihedral, as W_i is not induced. The other possibilities are A_4, S_4, and A_5, and of these only A_5 has elements of order 5. Thus $\overline{H}(i)$ is A_5, and H(i) is its double cover in $SL(W_i)$. So H(i) is abstractly the group $SL(2, \mathbb{F}_5)$, equipped with two noncommuting elements of order 5, A(i) and B(i). H(i) is then viewed as a subgroup of $SL(W_i)$ by a faithful irreducible two–dimensional representation of $SL(2, \mathbb{F}_5)$ which gives both A(i) and B(i) eigenvalues $\{\alpha. \beta\}$.

The group $SL(2, \mathbb{F}_5)$ has two inequivalent irreducible two–dimensional representations, say M_1 and M_2, which are $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$–conjugate. Both are faithful. In the group $SL(2, \mathbb{F}_5)$, the 24 elements of order five fall into two conjugacy classes, C_1 and C_2. Concretely C_1 is the conjugacy class of the upper unipotent matrix with 1 (or any nonzero square) in the upper corner, and C_2 is the conjugacy class of the upper unipotent matrix with 2 (or any nonzero nonsquare) in the upper corner. The classes C_1 and C_2 are interchanged by conjugation by any element in $GL(2, \mathbb{F}_5)$ with nonsquare determinant. Of the two representations M_i, one, say M_1, gives elements of C_1 eigenvalues $\{\alpha, \beta\}$ and elements of C_2 eigenvalues $\{\alpha^2, \beta^2\}$. The other, M_2, reverses this assignment. Since A and B both get eigenvalues $\{\alpha, \beta\}$ in each W_i, we may describe W_i as follows. We first take a surjective homomorphism

$$\pi(i) : H \to SL(2, \mathbb{F}_5)$$

which maps A and B to noncommuting elements A(i) and B(i) in the conjugacy class C_1, and then we embed $SL(2, \mathbb{F}_5)$ in SL(2) by M_1. We may further normalize this description of W_i as follows.

We may move A(i) by $SL(2, \mathbb{F}_5)$–conjugacy to $\mathrm{Unip}_+(1)$, the upper unipotent with upper corner 1. Having fixed A(i) as $\mathrm{Unip}_+(1)$, we may conjugate B(i) by the centralizer of $\mathrm{Unip}_+(1)$, which is $\pm 1\mathrm{Unip}_+$, and get B(i) to be one of the lower unipotents $\mathrm{Unip}_-(1)$ or $\mathrm{Unip}_-(-1)$. [Of the 12 elements in C_1, exactly two, $\mathrm{Unip}_+(1)$ and $\mathrm{Unip}_+(-1)$, commute with $\mathrm{Unip}_+(1)$. The remaining 10 fall into two orbits under conjugation by $\pm 1\mathrm{Unip}_+$, one of which contains $\mathrm{Unip}_-(1)$ and the other $\mathrm{Unip}_-(-1)$.]

With this normalization, the homomorphism

$$\pi(i) : H \to SL(2, \mathbb{F}_5)$$

is one of two possible maps, call them $\pi(+)$ and $\pi(-)$. The map $\pi(+)$, if it exists, maps A to $\mathrm{Unip}_+(1)$ and B to $\mathrm{Unip}_-(1)$. The map $\pi(-)$, if it exists, maps A to $\mathrm{Unip}_+(1)$ and B to $\mathrm{Unip}_-(-1)$. Depending on the relations satisfied by A and B in H, one of these maps might not exist as a homomorphism from H to $SL(2, \mathbb{F}_5)$.

If among the $\pi(i)$ only one of $\pi(+)$ or $\pi(-)$ occurs, then every W_i is $M \circ \pi(1)$. Pick an element D in $SL(2, \mathbb{F}_5)$ of order 6 (i.e., of trace 1). Pick an element E in H with $\pi(1)(E) = D$. Replacing E by E^{25^k} for large enough k, we may assume further that E has order prime to 5. Look at the action of E on

$$V|H = (\oplus_i W_i) \oplus (\oplus_j \chi_j).$$

Since the χ_j have order dividing 5, each $\chi_j(E) = 1$. In each W_i, E acts as M(D). As M is faithful, M(D) has order 6, so its eigenvalues are the two primitive sixth roots of unity, ζ_6 and its inverse. Thus E acts on V as

$$\mathrm{Diag}(\zeta_6 \text{ repeated k times}, \zeta_6^{-1} \text{ repeated k times}) \oplus (a \times a \text{ identity}),$$

where k is the number of W_i and a is the number of χ_j occurring in $V|H$. But we must have a=0, otherwise this element E, viewed in G, violates Blichfeldt's 60^o theorem, since it would have all its eigenvalues within 60^o of one of its eigenvalues, namely 1.

In this case, we can continue the analysis. Since $V|H$ is k copies of $M \circ \pi(1)$ and is a faithful representation, we conclude that $\pi(1)$ is an isomorphism $H \cong SL(2, \mathbb{F}_5)$.

If among the $\pi(i)$ both $\pi(+)$ and $\pi(-)$ occur (we will see below that this case does not arise), then every W_i is $M \circ \pi(+)$ or $M \circ \pi(-)$, and both occur, say k_+ and k_- times respectively. Then

$$V|H = (k_+ \text{ copies of } M \circ \pi(+)) \oplus (k_- \text{ copies of } M \circ \pi(-)) \oplus_j \chi_j.$$

We claim that the map

$$\pi(+) \times \pi(-) : H \to SL(2, \mathbb{F}_5) \times SL(2, \mathbb{F}_5)$$

is surjective. It suffices to show it induces a surjection

$$\bar{\pi}(+) \times \bar{\pi}(-) : H \to PSL(2, \mathbb{F}_5) \times PSL(2, \mathbb{F}_5),$$

simply because no proper subgroup of $SL(2, \mathbb{F}_5) \times SL(2, \mathbb{F}_5)$ maps onto $PSL(2, \mathbb{F}_5) \times PSL(2, \mathbb{F}_5)$. By Goursat's lemma [Lang, Algebra, ex. 5 on page 75], any subgroup of a product of two simple groups which maps onto each factor is either the whole product or the graph of an isomorphism. We can rule out having the graph of an isomorphism, because by direct calculation $\bar{\pi}(+)(AB)$ has order 5, while $\bar{\pi}(-)(AB)$ has order 3.

Pick an element D in $SL(2, \mathbb{F}_5)$ of order 6, and then pick an element E in H which, under $\pi(+) \times \pi(-)$, maps to (D, D). As above, we may choose E to have order prime to 5. Exactly as above, E acts on every W_i as M(D), and each $\chi_j(E) = 1$. Thus E acts on V as

$$\mathrm{Diag}(\zeta_6 \text{ repeated k times}, \zeta_6^{-1} \text{ repeated k times}) \oplus (a \times a \text{ identity}),$$

and, exactly as above, we infer that a=0 by Blichfeldt's 60^o theorem.

In this case too, we can continue the analysis. Since $V|H$ is k_+ copies of $M \circ \pi(+)$ and k_- copies of $M \circ \pi(+)$, and is a faithful representation, we conclude that $\pi(+) \times \pi(-)$ is an isomorphism $H \cong SL(2, \mathbb{F}_5) \times SL(2, \mathbb{F}_5)$, under which A is the element $(\mathrm{Unip}_+(1), \mathrm{Unip}_+(1))$ and under which B is the element $(\mathrm{Unip}_-(1), \mathrm{Unip}_-(-1))$.

So in either case, $V|H$ is $\oplus_i W_i$. As A acts on each W_i with eigenvalues $\{\alpha, \beta\}$, we get $n(\alpha) = n(\beta)$, as required.

In fact, as David Wales pointed out to me, this second case, when $\pi(+)\times\pi(-)$ is an isomorphism $H \cong SL(2, \mathbb{F}_5)\times SL(2, \mathbb{F}_5)$, does not arise. For if we take D in $SL(2, \mathbb{F}_5)$ an element of order 6, then the element (D, id) in $H \cong SL(2, \mathbb{F}_5)\times SL(2, \mathbb{F}_5)$ would act on some of the W_i as $(\zeta_6, \zeta_6{}^{-1})$, and on others as the identity, contradicting Blichfeldt's 60^o theorem. QED

Remark AZ.5 The proof of AZ.4 shows that the subgroup H constructed there is isomorphic to $SL(2, \mathbb{F}_5)$. In his survey paper [Zal, 11.2 and its proof], Zalesskii shows in addition that G = H.

AZ.6 Some Conjectures

(AZ.6.1) We end this appendix with several versions of a conjecture about what happens with quadratic elements of order 3 or 4.

Most optimistic conjecture AZ.6.2 (this has now been proven by Wales [Wales]) Over $\mathbb{C}$, suppose G is a finite irreducible primitive subgroup of GL(V) which contains a quadratic element

$$\gamma := \mathrm{Diag}(\zeta, \zeta, ..., \zeta, 1, 1, ..., 1)$$

of drop r, $1 \le r < \dim(V)$. Suppose that ζ is a primitive n'th root of unity, with $n \ge 3$. Then $\dim(V) \le 4r$.

(AZ.6.2.1) By Blichfeldt's 60^o theorem [Blich–FCG, paragraph 70, Theorem 8, page 96], this situation cannot arise with $n \ge 6$, and Zalesskii's result AZ.1 takes care of the case n=5. For n =3 or n=4, only the cases of low r seem to be in the literature. For r=1, the case of pseudoreflections, we have Mitchell's theorem [Mit]: $\dim(V) \le 2$ if n=4, and $\dim(V) \le 4$ if n=3. For r=2, we have the Huffman and Wales results [Huf–Wa]: $\dim(V) \le 4$ if n=4, and $\dim(V) \le 8$ if n=3. So one could even speculate, on the basis of this fairly limited range of numerical data, that for n = 4, we have $\dim(V) \le 2r$.

Optimistic conjecture AZ.6.3 There exists an integer $A \ge 4$ with the following property. Over $\mathbb{C}$, suppose G is a finite irreducible primitive subgroup of GL(V) which contains a quadratic element

$$\gamma := \mathrm{Diag}(\zeta, \zeta, ..., \zeta, 1, 1, ..., 1)$$

of drop r, $1 \le r < \dim(V)$. Suppose that ζ is a primitive n'th root of unity, with $n \ge 3$. Then $\dim(V) \le Ar$.

(AZ.6.3.1) Exactly as in the proof of Zalesskii's result AZ.1, to prove either of these first two versions of the conjecture, it suffices to treat the case where in addition the group G is generated by all the G–conjugates of γ.

Less optimistic conjecture AZ.6.4 There exists a polynomial P(x) in $\mathbb{Z}[x]$ with the following property. Over $\mathbb{C}$, suppose G is a finite irreducible primitive subgroup of GL(V) which contains a quadratic element

$$\gamma := \mathrm{Diag}(\zeta, \zeta, ..., \zeta, 1, 1, ..., 1)$$

of drop r, $1 \le r < \dim(V)$. Suppose that ζ is a primitive n'th root of unity, with $n \ge 3$. Then $\dim(V) \le P(r)$.

Least optimistic conjecture AZ.6.5 There exists a sequence $\{a(r)\}_{r\ge1}$ of integers with the following property. Over $\mathbb{C}$, suppose G is a finite irreducible primitive subgroup of GL(V) which contains a quadratic element

$$\gamma := \mathrm{Diag}(\zeta, \zeta, ..., \zeta, 1, 1, ..., 1)$$

of drop r, $1 \le r < \dim(V)$. Suppose that ζ is a primitive n'th root of unity, with $n \ge 3$. Then $\dim(V) \le a(r)$.

Chapter 2: Lefschetz Pencils, Especially on Curves

2.0 Review of Lefschetz pencils [SGA 7, Exposé XVII]

(2.0.1) We work over an algebraically closed field k. Let X/k be a proper smooth connected k–scheme of dimension $n \geq 1$, and $\mathcal{L}$ on X a very ample invertible $\mathcal{O}_X$–module. We embed X in $\mathbb{P}(H^0(X, \mathcal{L}))$, the projective space of hyperplanes in $H^0(X, \mathcal{L})$, in the usual way: x in X(k) is mapped to the hyperplane in $H^0(X, \mathcal{L})$ consisting of those global sections of $\mathcal{L}$ which vanish at x. Equivalently, we give ourselves X as a closed subscheme of a projective space $\mathbb{P}$ in such a way that both the following conditions are satisfied:

(2.0.1.1) $\mathcal{L}$ is $\mathcal{O}_X(1)$:= the pullback to X of $\mathcal{O}_{\mathbb{P}}(1)$,

(2.0.1.2) the restriction map induces an isomorphism

$$H^0(\mathbb{P}, \mathcal{O}_{\mathbb{P}}(1)) \cong H^0(X, \mathcal{O}_X(1)) := H^0(X, \mathcal{L}).$$

(2.0.2) A nonzero global section H of $H^0(\mathbb{P}, \mathcal{O}_{\mathbb{P}}(1))$ defines a hyperplane H=0, or simply H if no ambiguity is likely, in $\mathbb{P}$. The closed subscheme of X defined as X∩H is called the corresponding hyperplane section of X: in terms of the same global section H viewed as a global section H_X of $H^0(X, \mathcal{L})$, the hyperplane section X∩H is just the locus of vanishing of H_X as section of $\mathcal{L}$.

(2.0.3) Attached to this data, we have the dual variety $X^\vee$ in the dual projective space $\mathbb{P}^\vee$: it is the subset of $\mathbb{P}^\vee$ consisting of those hyperplanes H=0 in $\mathbb{P}$ such that X∩H fails to be smooth. It is known (SGA 7, Exposé XVII, 3.1.4) that $X^\vee$ is closed and irreducible, of codimension at least one in $\mathbb{P}^\vee$. [One sees $X^\vee$ as the image by the second projection of the closed subscheme Z of $X \times \mathbb{P}^\vee$ consisting of those pairs (x, H) such that H is tangent to X at x. The key point is that Z viewed over X is the total space of a $\mathbb{P}^{r-1}$ bundle over X, its projective normal bundle $\mathbb{P}(N_{X/\mathbb{P}})$, r the codimension of X in $\mathbb{P}$. Thus Z is proper and smooth over k, and $\dim(Z) = \dim(\mathbb{P}^\vee) - 1$. We endow $X^\vee$ with the induced reduced structure.

(2.0.4) Recall that a k–point of a k–scheme Y of dimension n−1 is called an ordinary double point if the complete local ring of Y at y is isomorphic to $k[[x_1, ..., x_n]]/Q(x)$, where Q(x) is given by

if n=2k is even, $Q(x) = \sum_{i=1 \text{ to } k} x_i x_{i+k}$,

if n=2k+1 is odd, $Q(x) = (x_{2k+1})^2 + \sum_{i=1 \text{ to } k} x_i x_{i+k}$.

(2.0.5) We denote by $\mathrm{Good}(X^\vee) \subset X^\vee$ those hyperplanes H such that the singular locus Sing(X∩H) of X∩H is a single point, say x_0, and such that X∩H has an ordinary double point at x_0. One knows [SGA 7, Exposé XVII, 3.2] that $\mathrm{Good}(X^\vee)$ is open in $X^\vee$. We denote by $\mathrm{Bad}(X^\vee) \subset X^\vee$ the closed complement of $\mathrm{Good}(X^\vee)$.

(2.0.6) Since $X^\vee$ is closed and irreducible in $\mathbb{P}^\vee$ of codimension at least one, we have:

Lemma 2.0.7 Given X in $\mathbb{P}$ as in 2.0.1, if $\mathrm{Good}(X^\vee)$ is nonempty, or if $X^\vee$ has codimension ≥ 2 in

$\mathbb{P}^\vee$, then $\mathrm{Bad}(X^\vee)$ has codimension ≥ 2 in $\mathbb{P}^\vee$.

Lemma 2.0.8 Given X in $\mathbb{P}$ as in 2.0.1, if $\mathrm{Good}(X^\vee)$ is nonempty, then $X^\vee$ is a hypersurface in $\mathbb{P}^\vee$.

proof Denote by $U \subset \mathbb{P}(N_{X/\mathbb{P}})$ the inverse image of $\mathrm{Good}(X^\vee)$ in the projective normal bundle. Then U is a nonempty and hence dense open set in $\mathbb{P}(N_{X/\mathbb{P}})$, so $\dim(U) = \dim(\mathbb{P}^\vee) - 1$. The map $U \to \mathrm{Good}(X^\vee)$ is bijective on k−valued points, hence $\dim(U) = \dim(\mathrm{Good}(X^\vee))$. As $\mathrm{Good}(X^\vee)$ is a nonempty and hence dense open set of $X^\vee$, we have $\dim(X^\vee) = \dim(U) = \dim(\mathbb{P}^\vee) - 1$. QED

(2.0.9) Recall that a Lefschetz pencil of hyperplane sections of X is a line L in $\mathbb{P}^\vee$, say $(\lambda, \mu) \mapsto \lambda F = \mu G$, such that the following two conditions hold.
(2.0.9.1) The "axis of the pencil", namely the codimension two linear subspace $\Delta$of $\mathbb{P}$ which is the common intersection of any two distinct members of the pencil (so here Δ is $F \cap G$) is transverse to X, i.e., $X \cap \Delta$ is smooth of codimension two in X. [The axis Δ determines the pencil, as consisting of all the hyperplanes containing Δ.]
(2.0.9.2) There is a dense open set U in $\mathbb{P}^1$ such that for (λ, μ) in U, $X \cap (\lambda F = \mu G)$ is smooth, and for (λ, μ) not in U, $X \cap (\lambda F = \mu G)$ is smooth outside a single point, where it has an ordinary double point.

(2.0.10) Equivalently, the lines L in $\mathbb{P}^\vee$ which are Lefschetz pencils of hyperplane sections of X are precisely those lines which satisfy the following three conditions.
(2.0.10.1) The axis Δ of L is transverse to X.
(2.0.10.2) L is not entirely contained in the dual variety $X^\vee$.
(2.0.10.3) $L \cap \mathrm{Bad}(X^\vee)$ is empty.

Proposition 2.0.11 Given X in $\mathbb{P}$ as in 2.0.1 above, suppose $\mathrm{Bad}(X^\vee)$ has codimension ≥ 2 in $\mathbb{P}^\vee$. Then we have:
1) The lines L in $\mathbb{P}^\vee$ which are Lefschetz pencils of hyperplane sections of X form a nonvoid (and hence dense) open set in the Grassmannian $\mathrm{Gr}(1, \mathbb{P}^\vee)$ of all lines in $\mathbb{P}^\vee$.
2) Let H be hyperplane such that $X \cap H$ is smooth. In the Grassmannian $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ of all lines in $\mathbb{P}^\vee$ which pass through H, the Lefschetz pencils of hyperplane sections of X form a nonvoid (and hence dense) open set in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$.
3) Let H be hyperplane such that $X \cap H$ has isolated singularities. In the Grassmannian $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ of all lines in $\mathbb{P}^\vee$ which pass through H, there is a dense open set U such that any L in U has the following three properties:

3a) the axis Δ of L is transverse to X,

3b) L is not entirely contained in the dual variety $X^\vee$,

3c) $L\cap\mathrm{Bad}(X^\vee)$ is either empty, if H lies in $\mathrm{Good}(X^\vee)$, or $L\cap\mathrm{Bad}(X^\vee)$ consists of H alone, if H lies in $\mathrm{Bad}(X^\vee)$.

proof For 1), note that each of the conditions 2.0.10.1–3 separately defines a nonvoid (and hence dense) open set in the Grassmannian, cf. [SGA 7, Exposé XVII, proof of 3.2.1]. For 2), it suffices to show that the dense open sets of $\mathrm{Gr}(1, \mathbb{P}^\vee)$ defined by the conditions 2.0.10.1–3 separately each have nonvoid intersection with $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$. For 2.0.10.1, there exist hyperplanes G transverse to $X\cap H$, and for any such G the pencil $\lambda G = \mu H$ satisfies 1a). Condition 2.0.10.2 holds on all of $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$, since H does not lie in $X^\vee$. The lines through H which violate 2.0.10.3 are the image Z of the proper scheme $\mathrm{Bad}(X^\vee)$ under the map $F \mapsto$ the line joining F to H. Thus Z is closed, and it has dimension $\dim(Z) \le \dim(\mathrm{Bad}(X^\vee)) \le \dim(\mathbb{P}^\vee) - 2$, while $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ has dimension $\dim(\mathbb{P}^\vee) - 1$.

For 3), we argue as follows. Conditions 3a) and 3b) each define open sets in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$. To obtain an L in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ for which 3a) holds, it suffices to find a hyperplane G such that $X\cap H\cap G$ is smooth (then take for L the line joining H to G). Such a G exists because $X\cap H$ has only isolated singularities: take a G which passes through none of the singular points of $X\cap H$, and which does not lie in the closure in $\mathbb{P}^\vee$ of the dual variety of $(X\cap H)^{\mathrm{smooth}}$. To exhibit a line L through H which does not lie entirely in $X^\vee$, take a hyperplane F not in $X^\vee$, and take for L the line joining H to F.

We now turn to condition 3c). Suppose first that H lies in $\mathrm{Good}(X^\vee)$. Then 3c) also defines a dense open set in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$, which one sees exactly as one saw in proving 2) above.

It remains to consider condition 3c) in the case in which H lies in $\mathrm{Bad}(X^\vee)$. In this case, we claim that the set, call it $\mathcal{S}$, of lines L in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ for which $L\cap\mathrm{Bad}(X^\vee)$ consists of H alone, **contains** a dense open set. The excluded lines through H are the image in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ of the scheme $\mathrm{Bad}(X^\vee) - \{H\}$ under the map $F \mapsto$ the line joining F to H. This image need not be closed, but its closure Z has dimension $\le \dim(\mathrm{Bad}(X^\vee)) \le \dim(\mathbb{P}^\vee) - 2$, while $\dim(\mathrm{Gr}(1, \mathbb{P}^\vee)_H) = \dim(\mathbb{P}^\vee) - 1$. Thus $\mathcal{S}$ contains the dense open set $\mathrm{Gr}(1, \mathbb{P}^\vee)_H - Z$. QED

Remark 2.0.12 It is the case 2) which is most commonly given, cf. [SGA 7, Exposé XVII, 3.2.8]. However, for our applications, 3) will be equally useful.

Definition 2.0.13 Let H be hyperplane such that $X\cap H$ has at worst isolated singularities. By a pencil through H which is **Lefschetz outside of H** we mean a line L through H which satisfies 3a), 3b), and 3c) of 2.0.11.

(2.0.14) In general, if we are given a pencil $(\lambda, \mu) \mapsto \lambda F = \mu G$ of hyperplanes in $\mathbb{P}$ whose axis is transverse to X, we form the incidence variety X' := the closed subscheme of $X \times \mathbb{P}^1$ consisting of pairs $(x, (\lambda, \mu))$ such that $\lambda F(x) = \mu G(x)$, and map it to $\mathbb{P}^1$ by the second projection. Because Δ is transverse to X, X' is smooth, being the blowup of X along the smooth subvariety $X \cap F \cap G$.

Theorem 2.0.15 Suppose that $\mathrm{Bad}(X^\vee)$ has codimension ≥ 2 in $\mathbb{P}^\vee$, and suppose that for every k–valued point x in X, we have $X^\vee \neq \mathrm{Hyp}_x$, the hyperplane in $\mathbb{P}^\vee$ consisting of all hyperplanes through x. Suppose we are given a hyperplane H such that $X \cap H$ has at worst isolated singularities. Suppose further that we are given a finite set S of k–valued points of X, none of which lies in $X \cap H$. Then in the Grassmannian $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ of all lines through H, there is a dense open set U such that every line L in U satisfies the following two conditions:
1) The pencil defined by L is Lefschetz outside of H.
2) Consider the map $f : X' \to \mathbb{P}^1$ defined by the pencil. View S as lying in X', by viewing $X - X \cap H$ as lying in X'. Then the points s in S lie in distinct fibres of the map $f : X' \to \mathbb{P}^1$, and each of these fibres $f^{-1}(f(s))$ is smooth.

proof Intrinsically, we may view the map $f : X' \to \mathbb{P}^1$ as having target the line L: for a point x in $X - X \cap \Delta$, $f(x) \in L$ is the unique point of intersection of L with the hyperplane Hyp_x in $\mathbb{P}^\vee$ of all hyperplanes through x. We already know that there is a dense open set U_1 in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ such that every line in U_1 satisfies 1). We will show that there exists a dense open set U_2 in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ such that every line in U_1 satisfies 2). Then the required U will be $U_1 \cap U_2$.

For each point s in S, we have $X^\vee \neq \mathrm{Hyp}_s$, hence $X^\vee \cap \mathrm{Hyp}_s$ has codimension at least two in $\mathbb{P}^\vee$. The hyperplanes $\{\mathrm{Hyp}_s\}_{s \text{ in } S}$ in $\mathbb{P}^\vee$ are all distinct, simply because $s \mapsto \mathrm{Hyp}_s$ is the canonical bijection {points in $\mathbb{P}$} $\cong$ {hyperplanes in $\mathbb{P}^\vee$). So for each pair s_i, s_j of distinct points of S, the intersection $\mathrm{Hyp}_{s_i} \cap \mathrm{Hyp}_{s_j}$ has codimension two in $\mathbb{P}^\vee$. The desired dense open set U_2 in $\mathrm{Gr}(1, \mathbb{P}^\vee)_H$ consists of those lines L through H which do not intersect the closed set

$$Z := \bigcup_{s \text{ in } S} \{X^\vee \cap \mathrm{Hyp}_s\} \bigcup_{i \neq j} \mathrm{Hyp}_{s_i} \cap \mathrm{Hyp}_{s_j}$$

in $\mathbb{P}^\vee$. The key point is that Z is a closed set of codimension at least two in $\mathbb{P}^\vee$, and Z does not contain H (since H contains none of the points s in S). The set U_2 is open by [EGA IV, Part 3, 13.1.5]. It is nonempty because if not, every line through H meets Z, and hence the map

$$Z \to Gr(1, \mathbb{P}^\vee)_H,\ z \mapsto \text{the line joining H to z}$$

is surjective, which is impossible since $\dim(Z) < \dim(Gr(1, \mathbb{P}^\vee)_H)$. QED

2.1 The dual variety in the favorable case

(2.1.1) We have the following basic result.

Proposition 2.1.2 [SGA 7, Exposé XVII, 3.3, 3.5] Given X in $\mathbb{P}$ as above, suppose that either dim(X) is even, or that char(k) $\neq$ 2. Suppose further that there exists a k–valued point x of X, and a hyperplane H such that $X \cap H$ contains x, and such that $X \cap H$ has an ordinary double point at x. Then $X^\vee$ is an irreducible hypersurface in $\mathbb{P}^\vee$, and $\mathrm{Good}(X^\vee)$ is its smooth locus $(X^\vee)^{\mathrm{smooth}}$.

Corollary 2.1.3 Hypotheses as in Proposition 2.1.2, $\mathrm{Bad}(X^\vee)$ is the singular locus $\mathrm{Sing}(X^\vee)$, and hence $\mathrm{Bad}(X^\vee)$ has codimension ≥ 2 in $\mathbb{P}^\vee$.

Lemma 2.1.4 (compare [Ka–Spacefill, Lemma 12]) Hypotheses as in Proposition 2.1.2 above, given a k–valued point x of X, there exists a hyperplane H which contains x and for which $X \cap H$ is smooth.

proof Given x, denote by $\mathrm{Hyp}_x \subset \mathbb{P}^\vee$ the hyperplane consisting of all hyperplanes H in $\mathbb{P}$ which contain x. Those H in Hyp_x for which $X \cap H$ is smooth form an open set U in Hyp_x. We must show that U is nonempty. If not, then we have an inclusion $\mathrm{Hyp}_x \subset X^\vee$. Since $X^\vee$ is an irreducible hypersurface in $\mathbb{P}^\vee$, we must have $\mathrm{Hyp}_x = X^\vee$. Then $X^\vee$ is smooth, and hence, by [SGA 7, Exposé XVII, 3.3, 3.5], the map from the projective normal bundle $\mathbb{P}(N_{X/\mathbb{P}})$ to $X^\vee$ is an isomorphism. Thus $\mathrm{Hyp}_x = X^\vee \cong \mathbb{P}(N_{X/\mathbb{P}})$ is a projective bundle over X, with fibre $\mathbb{P}^{r-1}$, r being the codimension of X in $\mathbb{P}$. [The careful reader at this point will ask what happens if r=0, i.e., if X is $\mathbb{P}$ itself. But this case is ruled out by the hypothesis of the Proposition that X has a hyperplane section which has an ordinary double point somewhere: if X were $\mathbb{P}$, every hyperplane section would be smooth.]

If r = 1, then $X^\vee \cong X$, and so X is isomorphic to a hyperplane. But X, being smooth of codimension one in $\mathbb{P}$, is a smooth hypersurface in $\mathbb{P}$, say of degree d. The degree d cannot be one, because we have assumed that X is embedded in $\mathbb{P}(H^0(X, \mathcal{L}))$, which for X a hyperplane would require taking the ambient space to be X itself, i.e., we would in fact have r = 0. If d $\geq$ 3, or if d=2 and dim(X) is even, then X is not isomorphic to a hyperplane, because its middle Betti number (say with $\overline{\mathbb{Q}}_\ell$ coefficients, ℓ any prime invertible in k) exceeds that of a hyperplane. If dim(X) $\geq$ 3 and d $\geq$ 2, again X is not isomorphic to a hyperplane, because, as Ofer Gabber explained to me, its degree d is an intrinsic invariant. Namely, for X a smooth hypersurface in $\mathbb{P}$ of dimension $n \geq 3$, Pic(X) is $\mathbb{Z}$, with a unique generator L which is ample, namely the restriction to X of $\mathcal{O}_{\mathbb{P}}(1)$. The dim(X)–fold self–intersection L^n of the unique ample generator is d, the degree of X in $\mathbb{P}$.

For r=1, this leaves only the case when d=2 and dim(X)=1, a case in which X **is** isomorphic to a hyperplane. The characteristic is not 2 and the field k is algebraically closed, so our smooth quadric X is given, in suitable projective coordinates in the ambient $\mathbb{P} = \mathbb{P}^2$, by the equation

$$\Sigma_{i=0 \text{ to } 2} (X_i)^2 = 0.$$

We will see directly from the equation that we do not have $X^\vee \subset \mathrm{Hyp}_x$ for any point x in X, indeed we do not have $X^\vee \subset \mathrm{Hyp}_x$ for any point x in $\mathbb{P}^2$. At the point (1, i, 0) of X, the tangent hyperplane to X has equation

$$X_0 + iX_1 = 0.$$

At the point (1, –i, 0), the tangent hyperplane has equation

$$X_0 - iX_1 = 0.$$

So any point on both these tangent hyperplanes has $X_0 = X_1 = 0$. Repeating this argument with the points (1, 0, ±i), we see that any point on all tangent hyperplanes has $X_0 = X_1 = X_2 = 0$, but there is no such point in $\mathbb{P}$. This concludes the proof in the r=1 case.

So suppose now that $r \geq 2$, pick a prime number ℓ invertible in k, and consider the Leray spectral sequence for the projective bundle π

$$\mathrm{Hyp}_X = X^\vee \cong \mathbb{P}(N_{X/\mathbb{P}}) \xrightarrow{\pi} X.$$

We first remark that X must be simply connected. Indeed, $\mathbb{P}^{r-1}$ is simply connected, so the projection π, being a Zariski–locally trivial $\mathbb{P}^{r-1}$ bundle, induces an isomorphism on fundamental groups: as the total space Hyp_X is itself simply connected, we infer that X is simply connected. Therefore the lisse sheaves $R^i\pi_*\mathbb{Q}_\ell$ on X are all constant, with value

$$R^i\pi_*\mathbb{Q}_\ell \cong H^i(\mathbb{P}^{r-1}, \mathbb{Q}_\ell)_X \quad \cong 0 \text{ if i odd or } i > 2r-2$$
$$\cong \mathbb{Q}_\ell(-i/2) \text{ if i even in } [0, 2r-2].$$

Therefore the Leray spectral sequence has E_2 terms given by

$$E_2{}^{p,q} = H^p(X, R^q\pi_*\mathbb{Q}_\ell) = H^p(X, \mathbb{Q}_\ell)\otimes H^q(\mathbb{P}^{r-1}, \mathbb{Q}_\ell),$$

and it abuts to $H^{p+q}(\mathrm{Hyp}_X, \mathbb{Q}_\ell)$. Now because π is projective and smooth, the pullback map $\pi^*: H^p(X, \mathbb{Q}_\ell) \to H^p(\mathrm{Hyp}_X, \mathbb{Q}_\ell)$ is injective. Therefore $H^p(X, \mathbb{Q}_\ell)$ vanishes unless p is even. From the formula for E_2, we see in turn that $E_2{}^{p,q}$ vanishes unless both p and q are even. As the differential d_r has bidegree (r, 1–r), it follows that the spectral sequence must degenerate at E_2. [Alternatively, we could appeal to the general result that Leray degenerates at E_2 for any proper smooth map with a proper smooth base, by a reduction to the case when k is the algebraic closure of a finite field. One then uses the fact that, by Deligne's Weil II, $E_2^{p,q}$ is pure of weight p+q, and d_r, being galois–equivariant, respects weight, so being of bidegree (r, 1–r) must vanish.]

From the degeneration, applied with p+q=2, we get

$$1 = b^2(\mathrm{Hyp}_x) = \dim E_2^{2,0} + \dim E_2^{0,2} = b^2(X) + b^2(\mathbb{P}^{r-1}) = b^2(X) + 1,$$

and thus $b^2(X) = 0$. This is impossible for a projective smooth connected X/k of dimension $n \geq 1$, since the class of a hyperplane is a nonzero element of $H^2(X, \mathbb{Q}_\ell(1))$. QED

Remark 2.1.5 If dim(X) = n is odd ≥ 3, a much shorter proof of Lemma 2.1.4 is to observe that the degree of the dual variety $X^\vee$ is **even** (and hence $X^\vee$ is not isomorphic to a hyperplane). Indeed, by [SGA7, Exposé XVIII, 3.2], the degree of the dual variety $X^\vee$ is equal to

$$(-1)^n(\chi(X) + \chi(X\cap\Delta) - 2\chi(X\cap(\text{general hyperplane } H))).$$

If X is odd–dimensional, so is $X\cap\Delta$, and hence both $\chi(X)$ and $\chi(X\cap\Delta)$ are even. We do not know an analogous shorter argument for X of even dimension.

(2.1.6) We should also point out that in characteristic two, there are smooth X's of every odd dimension whose dual variety **is** a hyperplane. Namely, in $\mathbb{P}^{2n}$, the variety X of equation

$$(X_0)^2 = \Sigma_{i=1 \text{ to } n} X_i X_{n+i}$$

has dual variety $X^\vee$ the hyperplane in the dual projective space consisting of all linear forms

$$\Sigma_{i=0 \text{ to } 2n} a_i X_i \text{ with } a_0 = 0.$$

So in this example, $X^\vee$ is Hyp_z for the point z = (1, 0, 0, 0, ..., 0) in $\mathbb{P}$, but the point z does not lie in X.

(2.1.7) Here is one criterion which insures that $X^\vee$ is not contained in Hyp_x for any k–valued point x in X. It will be used in the later discussion of Lefschetz pencils on curves, see 2.3.4.

Lemma 2.1.8 Given X in $\mathbb{P}$ as in 2.0.1, suppose that for any k–valued point x of X, there exists a k–valued point y of X, and a hyperplane H in $\mathbb{P}$, such that $X\cap H$ is singular at y, and such that $X\cap H$ does not contain x. Then $X^\vee$ is not contained in Hyp_x for any k–valued point x in X.

proof This is a tautology. QED

(2.1.9) In the rest of this chapter, we will study the case when X is a curve, and $\lambda F = \mu G$ is a pencil on X whose axis Δ is transverse to X. In this case, $X\cap\Delta$ will be empty, X' will be X, and the mapping of X' = X to $\mathbb{P}^1$ defined by the pencil is $x \mapsto (G(x), F(x))$, or more simply the rational function G/F.

2.2 Lefschetz pencils on curves in characteristic not 2

(2.2.1) In this section, we work over an algebraically closed field k in which 2 is invertible, and we take C/k a proper, smooth, connected curve, whose genus we denote g. Any effective divisor D on C of degree ≥ 2g+1 is very ample, i.e., the invertible sheaf $\mathcal{L}(D)$:= the inverse ideal sheaf $I(D)^{-1}$ is very ample, cf. [Hart, IV, 3.2 (b)].

Lemma 2.2.2 Fix an effective divisor D on C with deg(D) ≥ 2g+2, and use it to embed C in $\mathbb{P}$. For every k−valued point P on C, there exists a hyperplane H in $\mathbb{P}$ such that C∩H has an ordinary double point at P.

proof In the embedding by L(D) := $H^0(C, \mathcal{L}(D))$, a hyperplane section C∩H of C is the zero set of a nonzero element of L(D) (zero set **as** section of $\mathcal{L}(D)$). A hyperplane H such that C∩H has an ordinary double point at P is precisely the zero−locus on C of a nonzero element f of L(D) := $H^0(C, \mathcal{L}(D))$ which, as section of $\mathcal{L}(D)$, has a double zero at P. To see that such f exist, notice that the elements of L(D) with at least a double zero at P form the subspace L(D − 2P) of L(D), while those with at least a triple zero at P form the subspace L(D − 3P). Because deg(D) ≥ 2g+2, both D−2P and D−3P have degree ≥ 2g−1, so by Riemann−Roch we have

$$\ell(D-2P) = \deg(D-2P) + 1 - g = \deg(D) -1 -g,$$
$$\ell(D-3P) = \deg(D-3P) + 1 - g = \deg(D) -2 -g.$$

Therefore L(D − 3P) is a hyperplane in L(D − 2P), and any element of L(D − 2P) − L(D − 3P) is an f with a double zero (as section of $\mathcal{L}(D)$) at P. QED

(2.2.3) For degree 2g+1, we have:

Lemma 2.2.4 Suppose that C has genus g ≥ 1. Fix an effective divisor D on C with deg(D) = 2g+1, and use it to embed C in $\mathbb{P}$. For all but at most finitely many k−valued points P on C, there exists a hyperplane H in $\mathbb{P}$ such that C∩H has an ordinary double point at P.

proof Exactly as above, what we must prove is that for most points P in C(k), we have

$$\ell(D-2P) > \ell(D-3P).$$

Since deg(D−2P) = 2g−1 > 2g−2, we have

$$\ell(D-2P) = \deg(D-2P) + 1 - g = \deg(D) -1 -g.$$

But D−3P has degree 2g−2, so

$$\ell(D-3P) = \deg(D-3P) + 1 - g + \ell(K -(D-3P))$$
$$= \deg(D) -2 -g + \ell(K + 3P -D).$$

We must show that $\ell(K + 3P - D) = 0$ for most P. Since K + 3P −D has degree zero, we have

$$\ell(K + 3P -D) > 0 \text{ if and only if } K + 3P -D \text{ is a principal divisor.}$$

Consider the map $C \to \mathrm{Jac}^0(C)$ defined by

$$P \mapsto \text{the class of } K + 3P -D.$$

We claim this map has finite fibres (in which case only the finitely many P which map to the origin have $\ell(K + 3P - D) > 0$, and we are done). If not, then some fibre is infinite, and hence is all of C, i.e., the map is constant, which means in turn that for any two points P and Q in C(k), we have $3(P-Q) = 0$ in $\mathrm{Jac}^0(C)$. Fix Q. The map

$$P \mapsto P-Q$$

is a map from $C \mapsto \mathrm{Jac}^0(C)$ which lands in the finite set of points of order 3, hence is constant, hence (evaluate at P) has value 0, i.e., we find that the divisor P–Q is principal, say P–Q = div(f), in which case f is an isomorphism from C to $\mathbb{P}^1$, which is impossible since $g \geq 1$. QED

(2.2.5) In view of these lemmas 2.2.2 and 2.2.4, all the hypotheses of Proposition 2.1.2 of the previous section are satisfied, if $\deg(D) \geq \mathrm{Max}(2g+1, 2)$. Hence Theorem 2.0.15 of the last section holds. We apply it in the following way. We begin with our effective divisor D of degree $\geq \mathrm{Max}(2g+1, 2)$. We take for H the hyperplane defined by the vanishing of the section 1 of $I^{-1}(D)$, so $C \cap H$ is just D itself. To specify a pencil which passes through H and whose axis is transverse to C (i.e., whose axis is empty) is to give a second function f in $L(D) := H^0(C, I^{-1}(D))$ whose divisor of poles is precisely D (i.e., whose zeroes, as section of $I^{-1}(D)$, are disjoint from D). The resulting map of C to $\mathbb{P}^1$ is given by the ratio f/1 of these sections, i.e., it is given by f viewed as a rational function on C.

Theorem 2.2.6 Let k be an algebraically closed field in which 2 is invertible, and let C/k be a projective, smooth connected curve, of genus denoted g. Fix an effective divisor D on C of degree $d \geq 2g+1$. Fix a finite subset S of C – D. Then in L(D) viewed as the k–points of an affine space of dimension d+1–g, there is a dense open set U such that any f in U has the following properties:
1) the divisor of poles of f is D, and f is Lefschetz on C–D, i.e., if we view f as a finite flat map of degree d from C – D to $\mathbb{A}^1$, then the differential df on C–D has only simple zeroes, and f separates the zeroes of df (i.e., if α and β in C – D are zeroes of df, $f(\alpha) = f(\beta)$ if and only if $\alpha = \beta$. Put another way, all but finitely many of the fibres of f over $\mathbb{A}^1$ consist of d distinct points, and the remaining fibres consist of d–1 distinct points, d–2 of which occur with multiplicity 1, and one which occurs with multiplicity 2.
2) f separates the points of S, i.e., $f(s_1) = f(s_2)$ if and only if $s_1 = s_2$, and f is finite etale in a neighborhood of each fibre $f^{-1}(f(s))$. Put another way, there are #S fibres over $\mathbb{A}^1$ which each have d points and which each contain a single point of S.

proof If $\deg(D) \geq \mathrm{Max}(2g+1, 2)$, this is Theorem 2.0.15, specialized to curves. If g=0 and deg(D) = 1, then D is a single point, say ∞, C–D is $\mathbb{A}^1 := \mathrm{Spec}(k[x])$, L(D) is {1, x}, and the open set U consists of all functions ax+b with a, b in k and $a \neq 0$. QED

Remark 2.2.7 It is surely possible to prove this result entirely in the world of curves, but we believe that seeing it in the general context of Lefschetz pencils clarifies and simplifies what is going on. Caveat emptor.

Lemma 2.2.8 Hypotheses and notations as in Theorem 2.2.6 above, suppose the effective divisor D, which is the fibre of f over ∞ in $\mathbb{P}^1$, is $\sum a_i P_i$ with each a_i invertible in k. For f in the dense open set U, f viewed as map of C – D to $\mathbb{A}^1$ has $2g-2 + \sum (1 + a_i)$ singular fibres over $\mathbb{A}^1$, or, equivalently, df has $2g-2 + \sum (1 + a_i)$ zeroes.

proof Because each a_i is prime to p, df has a pole of order $1+ a_i$ at P_i. Since the canonical bundle has degree $2g-2$, the total number of zeroes of df, or what is the same, the number of singular fibres over $\mathbb{A}^1$, is $2g-2 + \sum(1 + a_i)$.QED

2.3 The situation for curves in arbitrary characteristic

(2.3.1) Let C/k be a proper smooth connected curve over an algebraically closed field k. Fix an effective divisor D of degree $d \geq 2g+3$, and use $\mathcal{L}(D)$ to embed C in $\mathbb{P}$.

Lemma 2.3.2 Let C/k be as in 2.3.1 above. Suppose $d \geq 2g+3$. For every k–valued point P on C, there exists a hyperplane H in $\mathbb{P}$ such that $C\cap H$ has an ordinary double point at P and such that $C\cap H$ is lisse outside of P. Moreover, the set of such H is an open dense set in the space of all hyperplanes tangent to C at P.

proof The hyperplanes H tangent to C at P are the points of the projective space $\mathbb{P}(L(D-2P)^\vee)$ of lines in $L(D-2P)$. In $\mathbb{P}(L(D-2P)^\vee)$, those for which $C\cap H$ does not have an ordinary double point at P are the points of the codimension one (by Riemann–Roch) subspace $\mathbb{P}(L(D-3P)^\vee)$. In $\mathbb{P}(L(D-2P)^\vee)$, the hyperplanes H for which $C\cap H$ has a singularity at a point $Q \neq P$ are the points of the codimension two (by Riemann–Roch) subspace $\mathbb{P}(L(D-2P - 2Q)^\vee)$.

We claim that In $\mathbb{P}(L(D-2P)^\vee)$, the union $\mathcal{W}$ over all Q (including Q=P) of the subspaces $\mathbb{P}(L(D-2P - 2Q)^\vee)$ is closed of codimension at least one. To see this, notice that there is a vector bundle $\mathcal{B}\mathrm{itan}_P$ on C whose fibre over Q is $L(D-2P-2Q)$. [Start with the line bundle $\mathcal{L}(D-2P)$ on C, and on C×C form the line bundle

$$\mathcal{L}_0 := (\mathrm{pr}_1{}^*\mathcal{L}(D-2P))\otimes I(\Delta)^{\otimes 2},$$

which on C×Q is $\mathcal{L}(D-2P-2Q)$, a line bundle of degree $d-4 > 2g-2$. Then $R^1\mathrm{pr}_{2*}\mathcal{L}_0 = 0$, and $\mathrm{pr}_{2*}\mathcal{L}_0$ is the desired vector bundle $\mathcal{B}\mathrm{itan}_P$ on C, whose formation commutes with arbitrary change of base on C.] The total space of the associated projective bundle $\mathbb{P}(\mathcal{B}\mathrm{itan}_P{}^\vee)$ is the closed subscheme W of $C \times \mathbb{P}^\vee$ consisting of all pairs (Q, H) with H in $\mathbb{P}(L(D-2P-2Q)^\vee)$, and $\mathcal{W}$ is the image of W under the second projection. Since W is proper and smooth over k of dimension $= \dim \mathbb{P}(L(D-2P\ \ )^\vee) - 1$, $\mathcal{W}$ is closed of codimension at least one in $\mathbb{P}(L(D-2P)^\vee)$.

Thus the set of hyperplanes H in $\mathbb{P}$ such that $C \cap H$ has an ordinary double point at P and such that $C \cap H$ is lisse outside of P are precisely the points of $\mathbb{P}(L(D-2P)^\vee)$ which do not lie in the proper closed subset $\mathcal{W} \cup \mathbb{P}(L(D-3P)^\vee)$. QED

Corollary 2.3.3 Suppose $d \geq 2g+3$. The dual variety $C^\vee$ has codimension one in $\mathbb{P}^\vee$. In $C^\vee$, the set $\mathrm{Good}(C^\vee)$ consisting of those hyperplanes H such that $C \cap H$ has just one singular point, and that one singular point is an ordinary double point, is a dense open set.

proof The dual variety $C^\vee$ has codimension at least one in $\mathbb{P}^\vee$. If the dual variety had codimension two or more in $\mathbb{P}^\vee$, we could find a Lefschetz pencil on C with no singular fibres (i.e., we could find a line L in $\mathbb{P}^\vee$ which did not meet $C^\vee$). The associated map to $\mathbb{P}^1$ would make C a finite etale connected covering of $\mathbb{P}^1$ of degree $d \geq 2g+3 > 1$, contradicting the fact that $\mathbb{P}^1$ is simply connected.

Once we know the dual variety is a hypersurface, it suffices to show that the hyperplanes H, such that $C \cap H$ has either two or more singularities, or has a singularity worse than an ordinary double point, form a closed set of codimension at least 2 in $\mathbb{P}^\vee$. Those with at least two singular points, or with one singularity which is a contact of order 4 or more, are the union $\mathcal{X}$ of the $\mathbb{P}(L(D-2P-2Q)^\vee)$ over all points (P, Q) in $C \times C$. Those with a singularity worse than on ordinary double point are the union $\mathcal{Y}$ of the $\mathbb{P}(L(D-3P)^\vee)$ over all points P in C.

We first deal with $\mathcal{X}$. On $C \times C$, there is a vector bundle $\mathcal{B}$itan whose fibre at (P, Q) is $L(D-2P-2Q)$. [Start with the line bundle $\mathcal{L}(D)$ on C, and on $C \times C \times C$ form the line bundle

$$\mathcal{L}_0 := (\mathrm{pr}_1{}^*\mathcal{L}(D)) \otimes I(\Delta_{1,2})^{\otimes 2} \otimes I(\Delta_{1,3})^{\otimes 2},$$

where $\Delta_{1,2}$ and $\Delta_{1,3}$ are the indicated partial diagonals. On $C \times P \times Q$, this line bundle is $\mathcal{L}(D-2P-2Q)$, a line bundle of degree $d-4 > 2g-2$. Then $R^1\mathrm{pr}_{2,3*}\mathcal{L}_0 = 0$, and $\mathrm{pr}_{2,3*}\mathcal{L}_0$ is the desired vector bundle $\mathcal{B}$itan on $C \times C$, whose formation commutes with arbitrary change of base on $C \times C$.] The total space of the associated projective bundle $\mathbb{P}(\mathcal{B}\mathrm{itan}^\vee)$ is the closed subscheme X of $C \times C \times \mathbb{P}^\vee$ consisting of all triples (P, Q, H) with H in $\mathbb{P}(L(D-2P-2Q)^\vee)$, and $\mathcal{X}$ is the image of X under the third projection. Since X is proper and smooth over k of dimension $\dim \mathbb{P}^\vee - 2$, $\mathcal{X}$ is closed of codimension at least two in $\mathbb{P}^\vee$.

We deal similarly with $\mathcal{Y}$. On C there is a vector bundle Triple whose fibre at P is $L(D-3P)$. The total space of the associated projective bundle $\mathbb{P}(\mathrm{Triple}^\vee)$ is the closed subscheme Y of $C \times \mathbb{P}^\vee$ consisting of all pairs (P, H) with H in $\mathbb{P}(L(D-3P)^\vee)$, and $\mathcal{Y}$ is the image of Y under the second projection. Since Y is proper and smooth over k of dimension $\dim \mathbb{P}^\vee - 2$, $\mathcal{Y}$ is closed of codimension at least two in $\mathbb{P}^\vee$. QED

Lemma 2.3.4 Suppose $d \geq 2g+3$. For every k−valued point P on C, and for every k−valued point $Q \neq P$ on C, there exists a hyperplane H in $\mathbb{P}$ such that $C \cap H$ is singular at Q, and such that $C \cap H$ does not contain P.

proof The hyperplanes H tangent to C at Q are the points of $\mathbb{P}(L(D-2Q)^\vee)$, a projective space of dimension $d-2-g \geq g+1$. Among all such H, those passing through P are in the subspace $\mathbb{P}(L(D-2Q-P)^\vee)$. As $d \geq 2g+2$, this is a subspace of codimension one. QED

Lemma 2.3.5 Suppose $d \geq 2g+3$. For every k−valued point P on C, there exists a hyperplane H through P such that $C \cap H$ is smooth.

proof Given P, denote by $\mathrm{Hyp}_P \subset \mathbb{P}^\vee$ the hyperplane consisting of all hyperplanes H in $\mathbb{P}$ which contain P. If no H in Hyp_P had $C \cap H$ smooth, we would have $\mathrm{Hyp}_P \subset C^\vee$. As $C^\vee$ is irreducible of codimension at most 1, this would force $\mathrm{Hyp}_P = C^\vee$, and this in turn would force $C^\vee \subset \mathrm{Hyp}_P$. But by the previous lemma, there are H in $C^\vee$ which do not contain P. QED

2.4 Lefschetz pencils on curves in characteristic 2

(2.4.1) We begin with the characteristic two version of Theorem 2.0.15.

Theorem 2.4.2 Let k be an algebraically closed field of characteristic 2, and let C/k be a projective, smooth connected curve, of genus denoted g. Fix an effective divisor D on C of degree $d \geq 2g+3$. Suppose that $D = \sum a_i P_i$. Fix a finite subset S of $C - D$. Then in L(D) viewed as the k−points of an affine space of dimension $d+1-g$, there is a dense open set U such that any f in U has the following properties:
1) the divisor of poles of f is D, and f is Lefschetz on $C-D$, i.e., if we view f as a finite flat map of degree d from $C - D$ to $\mathbb{A}^1$, then all but finitely many of the fibres of f over $\mathbb{A}^1$ consist of d distinct points, and the remaining fibres consist of $d-1$ distinct points, $d-2$ of which occur with multiplicity 1, and one which occurs with multiplicity 2.
2) f separates the points of S, i.e., $f(s_1) = f(s_2)$ if and only if $s_1 = s_2$, and f is finite etale in a neighborhood of each fibre $f^{-1}(f(s))$. Put another way, there are #S fibres over $\mathbb{A}^1$ which each have d points and which each contain a single point of S.

proof By Corollary 2.3.3 to Lemma 2.3.2 above, we know that $C^\vee$ is a hypersurface and that

Good($C^\vee$) is nonempty, and hence (by Lemma 2.0.7) that Bad($C^\vee$) has codimension at least two in $\mathbb{P}^\vee$. By Lemma 2.3.4 (and the tautologous Lemma 2.1.8), we know that $C^\vee$ is not contained in Hyp_P for any k–valued point P in C. Then by Theorem 2.0.15, we get a dense open set U_1 in L(D) such that every f in U_1 satisfies 1) and 2). QED

(2.4.3) The problem with this result is that it tells us nothing about the zeroes of the differential df of a function f in the open set U. This deficiency is remedied by the following theorem, which is the main result of this section.

Theorem 2.4.4 Let k be an algebraically closed field of characteristic 2, and let C/k be a projective, smooth connected curve, of genus denoted g. Fix an effective divisor D on C of degree $d \geq 6g+3$. Suppose that $D = \sum a_i P_i$ with each a_i odd. Fix a finite subset S of C – D. Then in L(D) viewed as the k–points of an affine space of dimension d+1–g, there is a dense open set U such that any f in U has the following properties:
1a) the divisor of poles of f is D, and f is Lefschetz on C–D, i.e., if we view f as a finite flat map of degree d from C – D to $\mathbb{A}^1$, then all but finitely many of the fibres of f over $\mathbb{A}^1$ consist of d distinct points, and the remaining fibres consist of d–1 distinct points, d–2 of which occur with multiplicity 1, and one which occurs with multiplicity 2.
1b) The differential df has $g-1 + \sum_i ((1+a_i)/2)$ distinct zeroes in C–D, and each zero is a double zero.
2) f separates the points of S, i.e., $f(s_1) = f(s_2)$ if and only if $s_1 = s_2$, and f is finite etale in a neighborhood of each fibre $f^{-1}(f(s))$. Put another way, there are #S fibres over $\mathbb{A}^1$ which each have d points and which each contain a single point of S.

2.5 Comments on Theorem 2.4.4
(2.5.1) Before giving the proof of the theorem, let us explain what problems we are fighting against in characteristic 2. In any other characteristic, once 1a) and 2) hold, then (as noted in Lemma 2.2.8 above) df has

$$2g-2 + \sum_i (1+a_i)$$

distinct zeroes, each of which is simple.
(2.5.2) The first problem is that in characteristic 2, for any function f on C, either df = 0, or df has all its zeroes and poles of **even** order. To see this, pick any k–valued point P on C, and any local parameter t at P, and expand f as a Laurent series in t, say

$$f = \sum b(n)t^n = \sum b(2n)t^{2n} + \sum b(2n+1)t^{2n+1}.$$

Because we are in characteristic 2, we get

$$df = \sum b(2n+1)t^{2n}dt.$$

(2.5.3) So we might hope that, if 1a) and 2) hold, then in characteristic two 1b) holds as well. But

1b) can fail spectacularly, even when 1a) and 2) hold.

(2.5.4) To illustrate most simply, consider the case when C is $\mathbb{P}^1$, and D is the divisor $(2k+1)\infty$, for some integer $k \geq 2$. The function $f(x) := x^2 + x^{2k+1}$ has divisor of poles D, and as a map of $C-D = \mathbb{A}^1$ to $\mathbb{A}^1$, f is Lefschetz. Indeed, there is only point x_0 at which df $(= x^{2k}dx)$ vanishes, namely $x_0 = 0$, and the fibre of f over the corresponding critical value $f(x_0) = 0$ is the zero set of

$$x^2 + x^{2k+1} = x^2(x^{2k-1} - 1),$$

which consists of 2k distinct points. But df has a single zero of order 2k, whereas 1b) calls for df to have $g-1 + (1+2k+1)/2 = k$ distinct zeroes, each of multiplicity 2.

2.6 Proof of Theorem 2.4.4

(2.6.1) By Theorem 2.4.2 above, we get a dense open set U_1 in L(D) such that every f in U_1 satisfies 1a) and 2).

(2.6.2) To complete the proof, it suffices to show that there is a dense open set U_2 in L(D) such that for f in U_2, f has polar divisor D and df has $g-1 + \Sigma_i ((1+a_i)/2)$ distinct zeroes in C−D, each a double zero. For then any f in the dense open set $U := U_1 \cap U_2$ will satisfy all of 1a), 1b), and 2).

Proposition 2.6.3 Let k be an algebraically closed field of characteristic 2, and let C/k be a projective, smooth connected curve, of genus denoted g. Fix an effective divisor $D = \Sigma a_i P_i$ on C of degree $d \geq 6g+3$. Suppose that each a_i is odd. Then in L(D) viewed as the k−points of an affine space of dimension d+1−g, there is a dense open set U_2 such that for f in U_2, f has polar divisor D and its differential df has $g-1 + \Sigma_i ((1+a_i)/2)$ distinct zeroes in C−D, each a double zero.

proof The proof is based upon the fact that in characteristic two, the canonical bundle $\Omega^1{}_{C/k}$ on a curve has a canonical square root, an observation that goes back to Mumford [Mum−TCAC]. Indeed, on an affine open piece Spec(A) of C which is etale over $\mathbb{A}^1{}_k :=$ Spec(k[x]) by a local coordinate x, the derivation d/dx on A has square zero, and both its kernel and its image consist precisely of the squares in A. In particular, for any f in A, df/dx is a square in A. So if we cover C by affine opens $\mathcal{U}_i :=$ Spec(A_i), each etale over $\mathbb{A}^1{}_k :=$ Spec(k[x_i]) by a local coordinate x_i, then $\Omega^1{}_{C/k}$ is locally free with basis dx_i on Spec(A_i). The transition functions $f_{i,j}$ defining $\Omega^1{}_{C/k}$ with respect to this covering are the ratios dx_i/dx_j on $\mathcal{U}_i \cap \mathcal{U}_j$. The key point is that these transition functions are **squares**, being of the form df/dx, and hence have unique square roots on $\mathcal{U}_i \cap \mathcal{U}_j$, say $f_{i,j} = (g_{i,j})^2$. The uniqueness guarantees that the $g_{i,j}$ form a 1−cocycle, and the line bundle $\mathcal{L}$ they define is the desired square root of the canonical bundle.

To put this into useful perspective, let us consider the more general situation of a smooth scheme X over a perfect field k of characteristic p > 0. We introduce the absolute Frobenius endomorphism $F : X \to X$, which on affine opens Spec(A) is $f \mapsto f^p$ on A. Then finding a p'th root of any line bundle on C amounts to descending it through F, i.e., writing it as $F^*(\mathcal{L})$ $(= \mathcal{L}^{\otimes p})$ for some line bundle $\mathcal{L}$ on C. Now there is a general result of Cartier, that to descend a quasicoherent sheaf $\mathcal{M}$ on X/k through the absolute Frobenius F is to give on $\mathcal{M}$ an integrable connection

$$\nabla : \mathcal{M} \to \mathcal{M} \otimes \Omega^1{}_{X/k}$$

of p–curvature zero, cf. [Ka–NCMT, 5.1].

Any connection is linear over the subsheaf of $\mathcal{O}_X$ consisting of p'th powers. Equivalently, if we take direct image by F, the connection map

$$\nabla : F_*\mathcal{M} \to F_*(\mathcal{M} \otimes \Omega^1{}_{X/k})$$

is $\mathcal{O}_X$–linear. Its kernel $\mathcal{N} := F_*\mathcal{M}^\nabla$ is thus a quasicoherent sheaf on X. Using the integrability and the fact that the p–curvature is zero, one shows that the canonical map $F^*\mathcal{N} \to \mathcal{M}$ is an isomorphism.

Let us return to our C/k of characteristic 2, and to the canonical square root $\mathcal{L}$ of the canonical bundle. The integrable connection of 2–curvature zero on $\Omega^1{}_{C/k}$ whose horizontal sections $(F_*\Omega^1{}_{C/k})^\nabla$ are $\mathcal{L}$ is precisely the integrable connection

$$\nabla : \Omega^1{}_{C/k} \to \Omega^1{}_{C/k} \otimes \Omega^1{}_{C/k}$$

given locally on Spec(A_i), A_i etale over $k[x_i]$, by defining ∇ to be the map $f dx_i \mapsto df \otimes dx_i$. This local description makes global sense precisely because the transition functions dx_j/dx_i are **squares**. The local horizontal sections are precisely (squares)dx_i, and these are in turn precisely the exact forms [simply because $f^2dx = d(f^2x)$]. More intrinsically, the local expression of the connection ∇ is

$$\nabla(fdg) := df \otimes dg.$$

Because the local horizontal sections of $F_*\Omega^1{}_{C/k}$ are the image of the exterior differentiation map

$$d: F_*\mathcal{O}_C \to F_*\Omega^1{}_{C/k},$$

we have a short exact sequence of locally free $\mathcal{O}_C$–modules

(2.6.3.1) $$0 \to \mathcal{O}_C \to F_*\mathcal{O}_C \to \mathcal{L} \to 0,$$

where the map $F_*\mathcal{O}_C \to \mathcal{L}$ is $f \mapsto \mathrm{Sqrt}(df)$.

Now take any divisor E on C, and tensor this short exact sequence with $I^{-1}(E)$. Since $F^*(I^{-1}(E)) = I^{-1}(2E)$, the middle term will be $I^{-1}(E) \otimes F_*\mathcal{O}_C \cong F_*F^*(I^{-1}(E)) = F_*(I^{-1}(2E))$, and we get

(2.6.3.2) $$0 \to I^{-1}(E) \to F_*(I^{-1}(2E)) \to \mathcal{L} \otimes I^{-1}(E) \to 0.$$

Here $\mathcal{L}\otimes I^{-1}(E)$ is the canonical descent of $I^{-1}(2E)\otimes\Omega^1{}_{C/k}$, and the map

$$F_*(I^{-1}(2E)) \to \mathcal{L}\otimes I^{-1}(E)$$

is $f \mapsto \mathrm{Sqrt}(df)$.

We now specialize this discussion to our effective divisor $D = \Sigma a_iP_i$ of degree $d \geq 6g+3$, all of whose coefficients a_i are odd. Since the a_i are all odd, exterior differentiation defines a map

$$F_*I^{-1}(\Sigma a_iP_i\ ) \to F_*(I^{-1}(\Sigma\ (a_i + 1)P_i)\otimes\Omega^1{}_{C/k}).$$

Because the a_i are odd, each $a_i + 1$ is even, and exterior differentiation **also** induces a map

$$F_*(I^{-1}(\Sigma\ (a_i + 1)P_i) \to F_*(I^{-1}(\Sigma\ (a_i + 1)P_i)\otimes\Omega^1{}_{C/k}).$$

This last map has precisely the same image as the one above, since we have only enlarged the source by allowing certain squares.

We have a short exact sequence

$$0 \to I^{-1}(\Sigma\ ((a_i + 1)/2)P_i) \to F_*(I^{-1}(\Sigma\ (a_i + 1)P_i)) \to I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L} \to 0,$$

which is just the exact sequence 2.6.3.2 above, with E taken to be the divisor

$$E = \Sigma\ ((a_i + 1)/2)P_i.$$

In view of the coincidence of images above, we also have a short exact sequence

$$0 \to I^{-1}(\Sigma\ ((a_i - 1)/2)P_i) \to F_*I^{-1}(\Sigma a_iP_i) \to I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L} \to 0.$$

The map

$$F_*I^{-1}(\Sigma a_iP_i) \to I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L}$$

is $f \mapsto \mathrm{Sqrt}(df)$. Its kernel consists of the squares in $F_*(I^{-1}(\Sigma a_iP_i))$, and these are precisely (remember each a_i is odd) the squares of local sections of $I^{-1}(\Sigma\ ((a_i - 1)/2)P_i)$.

In this context, we can now come to grips with showing that there is a dense open set U_2 of global sections of $F_*I^{-1}(\Sigma a_iP_i)$ for which df has precisely $g-1 + \Sigma((a_i + 1)/2)$ zeroes, each of which is a double zero. It is equivalent to show that there is a dense open set U_2 of global sections of $F_*I^{-1}(\Sigma a_iP_i)$ for which Sqrt(df) as global section of $I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L}$ has all its zeroes simple (the number of zeroes will then be $g-1 + \Sigma((a_i + 1)/2)$, which is the degree of $I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L}$.

As f runs over the global sections of $I^{-1}(\Sigma a_iP_i)$, the differentials df as global sections of $I^{-1}(\Sigma\ (a_i + 1)P_i)\otimes\Omega^1{}_{C/k}$ have no common zeroes. Indeed, by Theorem 2.4.2, part 1), a general global section f_1 of $I^{-1}(\Sigma a_iP_i)$ has exact divisor of poles Σa_iP_i, and hence df as section of

$$I^{-1}(\Sigma\ (a_i + 1)P_i)\otimes\Omega^1{}_{C/k}$$

is invertible near each P_i. But given any finite subset S of C – D, there is a dense open set of f's such that df is invertible near each s in S. Take S to be the zeroes of some df_1, and f_2 to have df_2 invertible both at the P_i and at the s in S. Then df_1 and df_2 have no common zeroes.

Therefore, as f runs over the global sections of $F_*I^{-1}(\Sigma a_iP_i)$, the global sections Sqrt(df) of $I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L}$ have no common zeroes. From the long exact cohomology sequence attached to the short exact sequence

$$0 \to I^{-1}(\Sigma\ ((a_i - 1)/2)P_i) \to F_*I^{-1}(\Sigma a_iP_i) \to I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L} \to 0,$$

we get a four term short exact sequence

$$0 \to H^0(C, I^{-1}(\Sigma\ ((a_i - 1)/2)P_i)) \to H^0(C, F_*I^{-1}(\Sigma a_iP_i))\to$$

$$\to H^0(C, I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L}) \to H^1(C, I^{-1}(\Sigma\ ((a_i - 1)/2)P_i)) \to 0.$$

The next term is

$$H^1(C, F_*I^{-1}(\Sigma a_iP_i)) \cong H^1(C, I^{-1}(\Sigma a_iP_i)) = 0,$$

the vanishing because Σa_iP_i has degree $\geq 6g+3 > 2g-2$. In our four–term exact sequence, we rewrite the second nonzero term:

$$H^0(C, F_*I^{-1}(\Sigma a_iP_i)) \cong H^0(C, I^{-1}(\Sigma a_iP_i)).$$

The first nonzero map

$$H^0(C, I^{-1}(\Sigma\ ((a_i - 1)/2)P_i)) \to H^0(C, I^{-1}(\Sigma a_iP_i))$$

is simply the squaring map, $f \mapsto f^2$. The second nonzero map is $f \mapsto$ Sqrt(df). The last term is $H^1(C, I^{-1}(\Sigma\ ((a_i - 1)/2)P_i))$, dual to a subspace of the holomorphic 1–forms, and so of dimension $\leq g$. [For example, if all $a_i = 1$, the last term will be $H^1(C, \mathcal{O})$.]

Thus our situation is the following. We have a line bundle

$$\mathcal{L}_1 := I^{-1}(\Sigma\ ((a_i + 1)/2)P_i)\otimes\mathcal{L},$$

whose degree is $\geq 4g+1$ (because $\geq (d+1)/2 + (g-1) \geq (6g+4)/2 + (g-1)$). Inside $H^0(C, \mathcal{L}_1)$ we have a linear subspace V, of codimension at most g, whose elements have no common zeroes (namely, the image of $H^0(C, I^{-1}(\Sigma a_iP_i))$ under the map $f \mapsto$ Sqrt(df)). We wish to show that for v in a dense open set $\mathcal{V}$ of V, v as section of $\mathcal{L}_1$ has all simple zeroes. (We then take U_2 to be the inverse image of $\mathcal{V}$ in $H^0(C, I^{-1}(\Sigma a_iP_i))$.) This results from the following elementary lemma.

Lemma 2.6.4 Let k be an algebraically closed field, C/k a proper smooth connected curve of genus g, $\mathcal{L}$ a line bundle of degree $d \geq 4g+1$, and $V \subset H^0(C, \mathcal{L})$ a linear subspace of codimension $\leq g$. Suppose that the elements of V have no common zeroes. Then the set $\mathcal{V} \subset V$ consisting of those v in V such that v as section of $\mathcal{L}_1$ has all simple zeroes is a dense open set of V.

proof First, let us remark that inside $\mathbb{P}(H^0(C, \mathcal{L})^\vee)$, the nonzero sections with all zeroes simple form a dense open set, say $\mathcal{U}$. [Its complement is the image of the total space of the projective bundle over C with fibre $\mathbb{P}(H^0(C, \mathcal{L}\otimes I(2P))^\vee)$ over the point P.] We must show that $\mathcal{V} := V\cap\mathcal{U}$ is nonempty.

Pick two nonzero elements v_0 and v_1 in $\mathcal{V}$ which have no common zero. Denote by D the divisor of zeroes of v_0. Then the map $f \mapsto fv_0$ is an isomorphism from $I^{-1}(D)$ to $\mathcal{L}$, which carries the global section 1 of $I^{-1}(D)$ to the global section v_0 of $\mathcal{L}$, and which carries some function f_1 in $H^0(C, I^{-1}(D))$ to the global section v_1. Because v_0 and v_1 have no common zeroes as sections of $\mathcal{L}$, the functions f_1 and 1 have no common zeroes as sections of $H^0(C, I^{-1}(D))$. More concretely, f_1 has its divisor of poles precisely equal to D.

Thus we are reduced to the case when $\mathcal{L}$ is $I^{-1}(D)$, with D an effective divisor of degree $d \geq 4g+1$, and when the linear subspace V of $H^0(C, I^{-1}(D))$ contains the function 1. Because $d \geq 2g$, the functions f in $H^0(C, I^{-1}(D))$ with exact divisor of poles D form a dense open set, say U. [The complement of U is the union, over the finitely many points P which occur in D, of the subspaces $H^0(C, I^{-1}(D-P))$, each of which has codimension 1 because $\deg(D) \geq 2g$.]

The open set $V\cap U$ of V is nonempty (it contains f_1), and hence is a dense open set of V.

If the ground field k has characteristic zero, pick any f in $V\cap U$. Then df is nonzero (because f is nonconstant), and hence has finitely many zeroes in C–D. Then for any λ in k which is not one of the finitely many critical values of f on C–D, the function $f - \lambda$ lies in V and has all its zeroes simple. Thus $f - \lambda$ lies in $\mathcal{V}$.

If the ground field k has characteristic $p > 0$, then we can repeat the same argument unless the f we choose in $V\cap U$ is a p'th power. Since f has divisor of poles D, f is a p'th power only if $D = pE$ for some (uniquely determined) effective divisor E, and f is g^p for some g in $H^0(C, I^{-1}(E))$.

If every f in $V\cap U$ is a p'th power, then

$$V\cap U \subset \text{p'th powers of elements of } H^0(C, I^{-1}(E)).$$

This leads to a contradiction, as follows. Comparing dimensions, we find

$$\dim(V) \leq \dim H^0(C, I^{-1}(E)).$$

A nonzero global section of $I^{-1}(E)$ has deg(E) zeroes, so we have the trivial inequality

$$\dim H^0(C, I^{-1}(E)) \leq 1 + \deg(E) = 1 + d/p.$$

On the other hand, V has codimension at most g in $H^0(C, I^{-1}(D))$, so

$$\dim(V) \geq d + 1 - g - g = d + 1 - 2g.$$

Thus we get the inequality

$$d + 1 - 2g \leq 1 + d/p,$$

or

$$d(p-1)/p \leq 2g,$$

i.e.,

$$d \le 2gp/(p-1) \le 4g,$$

contradiction. QED

2.7 Application to Swan conductors in characteristic 2

Theorem 2.7.1 Let k be an algebraically closed field of characteristic 2, and let C/k be a projective, smooth connected curve, of genus denoted g. Fix an effective divisor D on C of degree $d \ge 6g+3$. Suppose that $D = \sum a_iP_i$ with each a_i odd. Fix a finite subset S of C – D. Let f be any function in the open set U of Theorem 2.4.4. View f as a finite flat map of C–D to $\mathbb{A}^1$, and form the sheaf $\mathcal{F} := f_*\overline{\mathbb{Q}}_\ell$ on $\mathbb{A}^1$. Then $\mathcal{F}$ is tame at ∞. At each critical value α of f in $\mathbb{A}^1$, consider the $I(\alpha)$–representation $\mathcal{F}(\alpha)$. Then $I(\alpha)$ acts on $\mathcal{F}(\alpha)$ by a reflection of Swan conductor 1, i.e., $\mathcal{F}(\alpha)/\mathcal{F}(\alpha)^{I(\alpha)}$ is one–dimensional, and $I(\alpha)$ acts on $\mathcal{F}(\alpha)/\mathcal{F}(\alpha)^{I(\alpha)}$ by a character of order 2 having Swan conductor 1.

proof That $\mathcal{F}$ is tame at ∞ is immediate from the fact that f has a pole of order prime to the characteristic at each point of D. Because f is Lefschetz on C–D, for each critical value α of $\mathcal{F}$ in $\mathbb{A}^1$, $I(\alpha)$ acts on $\mathcal{F}(\alpha)$ by a reflection. The only question is to compute its Swan conductor. We have

$$\mathrm{Swan}_\alpha(\mathcal{F}(\alpha)) = \mathrm{Swan}_\alpha(\mathcal{F}(\alpha)/\mathcal{F}(\alpha)^{I(\alpha)}),$$

so what we must show is that each $\mathrm{Swan}_\alpha(\mathcal{F}(\alpha)) = 1$. Since the character $\mathcal{F}(\alpha)/\mathcal{F}(\alpha)^{I(\alpha)}$ has order 2 and we are in characteristic 2, we have an a priori inequality

$$\mathrm{Swan}_\alpha(\mathcal{F}(\alpha)) \ge 1.$$

Because df has $g-1 + \sum(1+a_i)/2$ zeroes, and f is Lefschetz, there are this many critical values. Thus it suffices to show that

$$\sum_{\alpha \text{ in } \mathrm{CritVal}(f)} (1 + \mathrm{Swan}_\alpha(\mathcal{F}(\alpha))) = 2g-2+\sum(1+a_i).$$

To show this, view $C - D - \cup_\alpha f^{-1}(\alpha)$ as a degree d finite etale covering of $\mathbb{A}^1 - \mathrm{CritVal}(f)$. Each fibre over a critical point α has d–1 instead of d points, so the Euler characteristic upstairs is given by

$$\chi(C - D - \cup_\alpha f^{-1}(\alpha), \overline{\mathbb{Q}}_\ell) = 2 - 2g - \#(D^{red}) - (d-1)\#\mathrm{CritVal}(f).$$

Computing downstairs, using the Euler–Poincaré formula and remembering that $\mathcal{F}$ is tame at ∞, we get

$$\chi(C - D - \cup_\alpha f^{-1}(\alpha), \overline{\mathbb{Q}}_\ell) = \chi(\mathbb{A}^1 - \mathrm{CritVal}(f), \mathcal{F})$$
$$= d(1 - \#\mathrm{CritVal}(f)) - \sum_{\alpha \text{ in } \mathrm{CritVal}(f)} \mathrm{Swan}_\alpha(\mathcal{F}(\alpha)).$$

Equating these two expressions for $\chi(C - D - \cup_\alpha f^{-1}(\alpha), \overline{\mathbb{Q}}_\ell)$, we get

$$2 - 2g - \#(D^{red}) - (d-1)\#\mathrm{CritVal}(f)$$
$$= d(1 - \#\mathrm{CritVal}(f)) - \Sigma_{\alpha \text{ in } \mathrm{CritVal}(f)} \mathrm{Swan}_{\alpha}(\mathcal{F}(\alpha)).$$

Cancelling the like term $-d\#\mathrm{CritVal}(f)$, we get

$$2 - 2g - \#(D^{red}) + \#\mathrm{CritVal}(f) = d - \Sigma_{\alpha \text{ in } \mathrm{CritVal}(f)} \mathrm{Swan}_{\alpha}(\mathcal{F}(\alpha)),$$

or, what is the same,

$$\Sigma_{\alpha \text{ in } \mathrm{CritVal}(f)} (1 + \mathrm{Swan}_{\alpha}(\mathcal{F}(\alpha))) = 2g-2 + \#(D^{red}) + d,$$

which is precisely the desired equality

$$\Sigma_{\alpha \text{ in } \mathrm{CritVal}(f)} (1 + \mathrm{Swan}_{\alpha}(\mathcal{F}(\alpha))) = 2g-2+\Sigma(1+a_i). \qquad \text{QED}$$

Remark 2.7.2 Suppose we take an f which satisfies conditions 1a) and 2) of Theorem 2.4.4, but not necessarily 1b). The above argument gives the equality

$$\Sigma_{\alpha \text{ in } \mathrm{CritVal}(f)} (1 + \mathrm{Swan}_{\alpha}(\mathcal{F}(\alpha))) = 2g-2 + \Sigma(1+a_i).$$

Therefore $\mathrm{Swan}_{\alpha}(\mathcal{F}(\alpha)) = 1$ for every critical point α if and only if f satisfies condition 1b) as well.

Chapter 3: Induction

3.0 The two sorts of induction

(3.0.1) Let G be a group, $H \subset G$ a subgroup, R a commutative ring, and V a left R[H]–module. There are two standard notions of the induction of V from H to G. The first, which we call "standard" induction, is

(3.0.1.1) $$\mathrm{Ind}_H^G(V) := R[G]\otimes_{R[H]}V,$$

with its structure of left R[G]–module through the first factor.
The second, which we call Mackey induction, is

(3.0.1.2) $$\mathrm{MaInd}_H^G(V) := \mathrm{Hom}_{\text{left } R[H]\text{–mod}}(R[G], V)$$
$$= \mathrm{Hom}_{\text{left } H\text{–sets}}(G, V),$$

which becomes a left R[G]–module by defining

$$(L_g\varphi)(x) := \varphi(xg).$$

(3.0.2) For standard induction, we get, for any left R[G]–module W, one version of Frobenius reciprocity:

(3.0.2.1) $$\mathrm{Hom}_{\text{left } R[H]\text{–mod}}(V, W|H)$$
$$\cong \mathrm{Hom}_{\text{left } R[G]\text{–mod}}(\mathrm{Ind}_H^G(V), W),$$

the isomorphism being $\psi \mapsto$ (the map $g\otimes v \mapsto g\psi(v)$). Taking for W the trivial R[G]–module R with trivial G–action, we get an isomorphism of coinvariants

(3.0.2.2) $$V_H \cong (\mathrm{Ind}_H^G(V))_G.$$

(3.0.3) For Mackey induction, we get the other version of Frobenius reciprocity:

(3.0.3.1) $$\mathrm{Hom}_{\text{left } R[H]\text{–mod}}(W|H, V)$$
$$\cong \mathrm{Hom}_{\text{left } R[G]\text{–mod}}(W, \mathrm{MaInd}_H^G(V)),$$

the isomorphism being $\psi \mapsto$ (the map $w \mapsto (g \mapsto \psi(gw))$). Taking for W the trivial R[G] module R with trivial G–action, we get an isomorphism of invariants

(3.0.3.2) $$V^H \cong (\mathrm{MaInd}_H^G(V))^G.$$

(3.0.4) When H has finite index in G, these two constructions are isomorphic, as follows. Define an R–linear map

(3.0.4.1) $$T : \mathrm{Hom}_{\text{left } H\text{–sets}}(G, V) \to R[G]\otimes_{R[H]}V$$

as follows. Pick any set of right coset representatives g_i for $H\backslash G$, i.e., G is the disjoint union of the right cosets Hg_i. Given an element φ in $\mathrm{Hom}_{\text{left } H\text{–sets}}(G, V)$, define $T(\varphi)$ to be the element

(3.0.4.2) $$T(\varphi) := \Sigma(g_i)^{-1}\otimes\varphi(g_i)$$

in $R[G]\otimes_{R[H]}V$. This map T is visibly an isomorphism of the underlying R–modules, each of which is #(G/H) copies of V.

(3.0.5) To see that T is well defined independent of the choice of right coset representatives g_i, notice that any other right coset representatives are of the form h_ig_i for some h_i in H. Then compute

$$\Sigma(h_ig_i)^{-1}\otimes\varphi(h_ig_i) = \Sigma(g_i)^{-1}(h_i)^{-1}\otimes h_i\varphi(g_i) = \Sigma(g_i)^{-1}\otimes\varphi(g_i).$$

To see that T is a homomorphism of left R[G]–modules, fix a in G, φ in $\mathrm{Hom}_{\text{left } H\text{–sets}}(G, V)$. Then

$$T(L_a(\varphi)) := \Sigma(g_i)^{-1}\otimes(L_a\varphi)(g_i) = \Sigma(g_i)^{-1}\otimes\varphi(g_ia)$$
$$= a(\Sigma a^{-1}(g_i)^{-1}\otimes\varphi(g_ia)) = a(\Sigma(g_ia)^{-1}\otimes\varphi(g_ia)) = a(T(\varphi)),$$

where we compute $T(\varphi)$ using the right coset representatives g_ia.

(3.0.6) If H is not assumed of finite index in G, then the above construction T establishes an isomorphism from the submodule of $\mathrm{Hom}_{\text{left } H\text{–sets}}(G, V)$ consisting of elements whose support is a finite union of right cosets of H in G, with $R[G]\otimes_{R[H]}V$.

(3.0.7) When H is of finite index in G, we write $\mathrm{Ind}_H{}^G(V)$ for "the" induction, and we have two Frobenius reciprocity isomorphisms:

(3.0.7.1) $$\mathrm{Hom}_{\text{left } R[H]\text{–mod}}(V, W|H) \cong \mathrm{Hom}_{\text{left } R[G]\text{–mod}}(\mathrm{Ind}_H{}^G(V), W),$$

and

(3.0.7.2) $$\mathrm{Hom}_{\text{left } R[H]\text{–mod}}(W|H, V) \cong \mathrm{Hom}_{\text{left } R[G]\text{–mod}}(W, \mathrm{Ind}_H{}^G(V)).$$

3.1 Induction and duality

(3.1.1) Let H be a group, R a commutative ring, and V a left R[H]–module whose underlying R–module is free of finite rank. Denote by $V^\vee$ the dual ("contragredient") representation. Its underlying R–module is

$$V^\vee := \mathrm{Hom}_{R\text{–mod}}(V, R),$$

and the left H–action on $V^\vee$ is defined as follows: given an R–linear map $\varphi : V \to R$, we define $h\varphi$ to be the R–linear map $v \mapsto \varphi(h^{-1}v)$. Thus the canonical pairing

$$< , > : V \times V^\vee \to R$$
$$<v, \varphi> := \varphi(v),$$

has the equivariance property that for all h in H, v in V, φ in $V^\vee$,

$$<hv, h\varphi> = <v, \varphi>.$$

(3.1.2) Equivalently, suppose we are given two left R[H]–modules V and W, both of whose underlying R–modules are free of finite rank, and an R–bilinear pairing

$$< , > : V \times W \to R$$

which is H–equivariant:

$$<hv, hw> = <v, w>.$$

If this pairing makes V and W R–duals of each other, then V and W are the contragredients of each other.

Lemma 3.1.3 Given V and W as in 3.1.2 above which are contragredients of each other, with pairing $< , >_H$, suppose that H is a subgroup of finite index in G. Then $\mathrm{Ind}_H{}^G(V)$ and $\mathrm{Ind}_H{}^G(W)$ are contragredients of each other:

$$\mathrm{Ind}_H{}^G(V)^\vee \cong \mathrm{Ind}_H{}^G(V^\vee).$$

proof The simplest way to see this is to think of induction as Mackey induction, and to write down $< , >_G$ a priori. To do this, pick a set of coset representatives g_i for H\G. Given maps of left H–sets

$$f_1 : G \to V \text{ and } f_2 : G \to W,$$

we define

$$< , >_G : \mathrm{Ind}_H{}^G(V) \times \mathrm{Ind}_H{}^G(W) \to R$$

by

$$<f_1, f_2>_G := \Sigma <f_1(g_i), f_2(g_i)>_H.$$

This pairing visibly makes the underlying R–modules R–duals of each other.

This pairing is independent of the choice of coset representatives g_i. Indeed, any other choice is $h_i g_i$ for some elements h_i in H, and for this new choice the individual summands remain unchanged:

$$<f_1(h_i g_i), f_2(h_i g_i)>_H = <hf_1(g_i), hf_2(g_i)>_H = <f_1(g_i), f_2(g_i)>_H.$$

The pairing thus defined is G–equivariant. For a in G,

$$<L_a f_1, L_a f_2>_G := \Sigma <(L_a f_1)(g_i), (L_a f_2)(g_i)>_H$$
$$= \Sigma <f_1(g_i a), f_2(g_i a)>_H.$$

This last sum is simply the expression of $<f_1, f_2>_G$ using the right coset representatives $g_i a$. QED

Corollary 3.1.4 Hypotheses and notations as in 3.1.3, if V is orthogonally (respectively symplectically) self–dual, then $\mathrm{Ind}_H{}^G(V)$ is orthogonally (respectively symplectically) self–dual.

proof If the form $< , >_H$ on V×V is symmetric (respectively strongly alternating, i.e., $<v,v>_H = 0$ for all v in V) then the bilinear form $<f_1, f_2>_G$ is symmetric (respectively strongly alternating). QED

3.2 Induction as direct image

(3.2.1) Suppose X and Y are connected schemes, and $f : X \to Y$ is a finite etale map. Then $H := \pi_1(X, \text{any base point } x)$ is an open subgroup of finite index in

$G := \pi_1(Y, \text{the base point } f(x))$. If R is a topological ring, for instance $\mathbb{F}_\ell$ or $\overline{\mathbb{F}}_\ell$ or $\mathbb{Z}_\ell$ or $\mathbb{Q}_\ell$ or $\overline{\mathbb{Q}}_\ell$, we may view a continuous representation V of H on, say, a free R−module of finite rank, as (the stalk at x of) a lisse sheaf $\mathcal{F}$ of R−modules on X. The direct image $f_*\mathcal{F}$ is the lisse sheaf of R−modules on Y corresponding to the induction of V from H to G. For $\mathcal{G}$ a lisse sheaf of R−modules on Y, corresponding to a continuous representation W of G, W|H corresponds to the lisse sheaf $f^*\mathcal{G}$ on X. So viewed, the second (and less standard) form of Frobenius reciprocity becomes the standard adjunction isomorphism

(3.2.1.1) $$\mathrm{Hom}_X(f^*\mathcal{G}, \mathcal{F}) \cong \mathrm{Hom}_Y(\mathcal{G}, f_*\mathcal{F}),$$

while the first (and more standard) form of Frobenius reciprocity becomes the exotic adjunction isomorphism

(3.2.1.2) $$\mathrm{Hom}_X(\mathcal{G}, f^!\mathcal{F}) \cong \mathrm{Hom}_Y(f_!\mathcal{G}, \mathcal{F}),$$

cf. [SGA 4, Exposé XVIII, 3,1,4,3].

3.3 A criterion for the irreducibility of a direct image

Proposition 3.3.1 (Irreducible Induction Criterion) Let k be an algebraically closed field, C_1/k and C_2/k two smooth connected curves, and $f: C_1 \to C_2$ a finite flat map of degree $d \geq 1$ which is generically etale. Let ℓ be a prime number invertible in k, and let $\mathcal{F}$ be an irreducible middle extension $\overline{\mathbb{Q}}_\ell$−sheaf on C_1, i.e., $\mathcal{F}$ is the extension by direct image of a lisse irreducible $\overline{\mathbb{Q}}_\ell$−sheaf on a dense open set of C_1. Suppose that $\mathrm{Sing}(\mathcal{F})$, the set of points at which $\mathcal{F}$ is not lisse, is nonempty. Suppose further that for some s in $\mathrm{Sing}(\mathcal{F})$, the fibre $f^{-1}(f(s))$ consists of d distinct points, only one of which lies in $\mathrm{Sing}(\mathcal{F})$. Then $f_*\mathcal{F}$ on C_2 is an irreducible middle extension.

proof We first recall why $f_*\mathcal{F}$ is a middle extension. Let U_2 in C_2 be a dense open set over which f is finite etale, and such that $f^{-1}(U_2)$ does not meet $\mathrm{Sing}(\mathcal{F})$ (i.e., such that $\mathcal{F}$ is lisse on $f^{-1}(U_2)$). Then we have a commutative diagram

$$\begin{array}{ccc} f^{-1}(U_2) & \xrightarrow{j_1} & C_1 \\ f_0 \downarrow & & \downarrow f \\ U_2 & \xrightarrow[j_2]{} & C_2 \end{array}.$$

Here $\mathcal{F}$ is $j_{1*}j_1{}^*\mathcal{F}$, so $f_*\mathcal{F}$ is $f_*j_{1*}j_1{}^*\mathcal{F} = j_{2*}f_{0*}j_1{}^*\mathcal{F}$. The sheaf $f_{0*}j_1{}^*\mathcal{F}$ on U_2 is lisse ($\mathcal{F}$ is lisse on $f^{-1}(U_2)$, and f_0 is finite etale), and it is equal to $j_2{}^*f_*\mathcal{F}$ (commutation of f_* with localization on the base). Thus $f_*\mathcal{F}$ is $j_{2*}j_2{}^*f_*\mathcal{F}$, as required.

It remains to prove that $f_*\mathcal{F}$ is irreducible on U_2. By assumption, $\mathcal{F}|f^{-1}(U_2)$ is a continuous irreducible $\overline{\mathbb{Q}}_\ell$–representation of $\pi_1(f^{-1}(U_2))$, an open subgroup of finite index d in $\pi_1(U_2)$. The lisse sheaf $(f_*\mathcal{F})|U_2$ is the induced representation of $\pi_1(U_2)$, and is therefore completely reducible (because we have coefficients $\overline{\mathbb{Q}}_\ell$ of characteristic zero: this complete reducibility can fail for $\overline{\mathbb{F}}_\ell$ coefficients, just think of taking $[\ell]_*\mathbb{F}_\ell$ for the ℓ'th power map of $\mathbb{G}_m$ to itself).

So $f_*\mathcal{F}$ on U_2 is irreducible if and only if $\mathrm{Hom}_{U_2}(f_*\mathcal{F}, f_*\mathcal{F})$ is one–dimensional, or, equivalently, has dimension < 2. By adjunction, we have

$$\mathrm{Hom}_{U_2}(f_*\mathcal{F}, f_*\mathcal{F}) = \mathrm{Hom}_{f^{-1}(U_2)}(f^*f_*\mathcal{F}, \mathcal{F}).$$

Once again, $f^*f_*\mathcal{F}$ is completely reducible on $f^{-1}(U_2)$, and the dimension of

$$\mathrm{Hom}_{f^{-1}(U_2)}(f^*f_*\mathcal{F}, \mathcal{F})$$

is the multiplicity of $\mathcal{F}$ in $f^*f_*\mathcal{F}$.

So what we must show is that $\mathcal{F}\oplus\mathcal{F}$ is not a direct summand of $f^*f_*\mathcal{F}$. To see this, we will show that already as representations of the inertia group I(s), $\mathcal{F}\oplus\mathcal{F}$ is not a direct summand of $f^*f_*\mathcal{F}$. Recall that f is etale at every point of the fibre $f^{-1}(f(s))$, and that $\mathcal{F}$ is lisse at every point of this fibre except for the point s itself. Therefore as a representation of I(s), $f^*f_*\mathcal{F}$ is the direct sum of $\mathcal{F}(s)$ and of d−1 trivial rank($\mathcal{F}$)–dimensional representations of I(s). Because $\mathcal{F}(s)$ is a nontrivial representation of I(s), we claim that $\mathcal{F}(s)\oplus\mathcal{F}(s)$ is not a direct summand of $\mathcal{F}(s)\oplus$(trivial). Indeed, if it were, then by Jordan Holder the semisimplification of $\mathcal{F}(s)$ would be trivial, i.e., $\mathcal{F}(s)$ would be a unipotent representation of I(s), i.e., a homomorphism from I(s) to the group of upper unipotent matrices. If $\mathcal{F}(s)$ is nontrivial, then some element γ of I(s) has a nontrivial Jordan normal form. From the theory of Jordan normal form we see that even after restriction to the cyclic subgroup $\langle\gamma\rangle$, $\mathcal{F}(s)\oplus\mathcal{F}(s)$ is not a direct summand of $\mathcal{F}(s)\oplus$(trivial). QED

Remark 3.3.2 If $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ is empty, $f_*\mathcal{F}$ need not be irreducible (e.g., for $\mathcal{F}$ the constant sheaf $\overline{\mathbb{Q}}_\ell$, $\overline{\mathbb{Q}}_\ell$ is always a direct factor of $f_*\overline{\mathbb{Q}}_\ell$).

3.4 Autoduality and induction

Proposition 3.4.1 Hypotheses and notations as in the Irreducible Induction Criterion 3.3.1 above, $\mathcal{F}$ on C_1 is self–dual if and only if $f_*\mathcal{F}$ on C_2 is self–dual. If both are self–dual, either they are both orthogonally self–dual, or they are both symplectically self–dual.

proof The implication $\Rightarrow$ is Corollary 3.1.4 above. For the converse, suppose that $f_*\mathcal{F}$ is self–dual, but that $\mathcal{F}$ is not self–dual. We arrive at a contradiction as follows. We know that $f_*\mathcal{F}$ is irreducible, and hence that $\mathcal{F}$ occurs in $f^*f_*\mathcal{F}$ as a direct summand. Because $f_*\mathcal{F}$ is self–dual, we have

$$f_*\mathcal{F} \cong (f_*\mathcal{F})^\vee \cong f_*(\mathcal{F}^\vee).$$

Therefore $\mathcal{F}^\vee$ occurs in $f^*f_*\mathcal{F}$ as a direct summand. If $\mathcal{F}$ is not isomorphic to $\mathcal{F}^\vee$, then $\mathcal{F}\oplus\mathcal{F}^\vee$ is a direct summand of $f^*f_*\mathcal{F}$. Looking at stalks at s, we get that $\mathcal{F}(s)\oplus\mathcal{F}(s)^\vee$ is a direct summand of $\mathcal{F}(s)\oplus$(trivial), which leads to a contradiction exactly as in the proof of 3.3.1 above.

Suppose now that $\mathcal{F}$ and $f_*\mathcal{F}$ are both self–dual. Since they are both irreducible, each admits a unique (up to a $\overline{\mathbb{Q}}_\ell^\times$–factor) autoduality, and that autoduality is either symplectic or orthogonal. By Corollary 3.1.4 above, once $\mathcal{F}$ is self–dual, either orthogonally or symplectically, $f_*\mathcal{F}$ is autodual of the same sort. QED

3.5 A criterion for being induced

(3.5.1) We work over an algebraically closed field K of characteristic zero. We are given a group G and a subgroup $H \subset G$ of finite index. Given a K–representation $\Lambda: H \to GL(W)$ of H, denote by

$$ch_\Lambda : H \to K$$

its character

$$ch_\Lambda(h) := \mathrm{Trace}(\Lambda(h)).$$

Extend the function ch_Λ by zero to all of G, i.e., consider the function

$$ch_{!\Lambda} : G \to K$$

defined by

$$ch_{!\Lambda}(g) = ch_\Lambda(g), \text{ if g lies in H,}$$
$$ch_{!\Lambda}(g) = 0 \text{ if g does not lie in H.}$$

Denote by $\mathrm{Ind}_H{}^G(\Lambda)$, or simply $\mathrm{Ind}(\Lambda)$, the G–representation $\mathrm{Ind}_H{}^G(W)$. One sees easily from the definitions that the character of $\mathrm{Ind}(\Lambda)$ is the function on G defined by

$$ch_{\mathrm{Ind}(\Lambda)}(g) := \textstyle\sum_{\gamma \text{ rep's of } G/H} ch_{!\Lambda}(\gamma h\gamma^{-1}).$$

Thus the character of Ind(H) is supported in $\cup_{g \text{ in } G}\, gHg^{-1}$.

(3.5.1.1) Suppose now in addition that H is a **normal** subgroup of G. Then the character of $\mathrm{Ind}(\Lambda)$ vanishes outside of H. To what extent is it true that an irreducible representation ρ of G whose character is supported in a normal subgroup $H \subset G$ of finite index is induced from H? Here is a very partial answer.

Theorem 3.5.2 Suppose $H \subset G$ is a normal subgroup of finite index, and that the quotient group G/H has squarefree order $N \geq 2$. Suppose given $\rho : G \to GL(V)$ an irreducible, finite–dimensional representation of G, whose character ch_ρ is supported in H. Then there exists an irreducible representation Λ of H such that $\rho \cong Ind_H{}^G(\Lambda)$. If in addition $\dim(\rho) = \#(G/H)$, then Λ is a (linear) character of H, i.e., $\dim(\Lambda) = 1$.

proof Because ρ is irreducible, it is completely reducible. Because H is normal in G (or because H is of finite index in G and char(K) = 0), $\rho|H$ is completely reducible, say

$$\rho|H = \Sigma_{i=1 \text{ to } r}\, n_i\Lambda_i.$$

Because H is normal in G, and ρ is irreducible on G, the various Λ_i are all G–conjugate, and all the n_i have a common value n:

$$\rho|H = \Sigma_{i=1 \text{ to } r}\, n\Lambda_i.$$

Recall that for any completely reducible finite–dimensional K–representation Λ of H, $Ind(\Lambda)$ is completely reducible on G. (Since K is of characteristic zero, it suffices to check complete reducibility of the restriction of $Ind(\Lambda)$ to any normal subgroup Γ in G of finite index; if we take Γ to be H itself, we are looking at $Ind(W)|H$, which is the direct sum $\oplus_{\gamma \text{ rep's of } G/H}\, \Lambda^{(\gamma)}$ of conjugates of Λ.)

For any two completely reducible finite–dimensional K–representations ρ and σ of G, we denote as usual

$$<\pi, \sigma>_G := \dim_K Hom_{K[G]-mod}(\pi, \sigma),$$

and similarly for H. Frobenius reciprocity now takes the following numerical form: for π (respectively τ) a completely reducible finite–dimensional K–representation of G (respectively of H), we have

$$<\tau, \pi|H>_H = <Ind(\tau), \pi>_G.$$

We now apply this to our situation. Recall that

$$\rho|H = \Sigma_{i=1 \text{ to } r}\, n\Lambda_i.$$

Thus

$$<\rho|H, \rho|H>_H := <\Sigma_{i=1 \text{ to } r}\, n\Lambda_i, \Sigma_{i=1 \text{ to } r}\, n\Lambda_i>_H = \Sigma_{i=1 \text{ to } r}\, n^2 = r\times n^2.$$

On the other hand, Frobenius reciprocity gives

$$<\rho|H, \rho|H>_H = <Ind_H{}^G(\rho|H), \rho>_G.$$

On the other hand, by the "projection formula", we have

$$Ind_H{}^G(\rho|H) \cong \rho\otimes_K Ind_H{}^G(\mathbb{1}_H).$$

Now $Ind_H{}^G(\mathbb{1}_H)$ is the regular representation of G/H, viewed as a representation of G. So its character vanishes outside of H, and is equal to #(G/H) at every h in H. Since the character of ρ is itself supported in H, we have

$$\mathrm{ch}_{\mathrm{Ind}(\rho|H)} = \#(G/H)\times \mathrm{ch}_{\rho} = \mathrm{ch}_{\#(G/H)\text{ copies of }\rho}.$$

Because completely reducible representations over an algebraically closed field of characteristic zero are determined up to isomorphism by their characters, we find

$$\mathrm{Ind}_H{}^G(\rho|H) \cong \#(G/H) \text{ copies of } \rho.$$

Returning to the inner products above, we get

$$<\rho|H, \rho|H>_H = <\mathrm{Ind}_H{}^G(\rho|H), \rho>_G = <\#(G/H) \text{ copies of } \rho, \rho>_G = \#(G/H).$$

Comparing the two evaluations of $<\rho|H, \rho|H>_H$, we find

$$r\times n^2 = \#(G/H).$$

But #(G/H) is squarefree, so we infer that n=1, r = #(G/H). Thus $\rho|H$ is the direct sum of #(G/H) distinct irreducibles Λ_i of H, which are transitively permuted by G–conjugation. This means precisely that H is the stabilizer of the isomorphism class of any single Λ_i, and that for each i we have

$$\rho \cong \mathrm{Ind}_H{}^G(\Lambda_i).$$

Once we know this, taking dimensions we get

$$\dim(\rho) = \#(G/H)\dim(\Lambda),$$

which makes obvious the final assertion of the theorem. QED

Remark 3.5.3 Suppose that the group G in Theorem 3.5.2 above is a topological group, H is an open and closed normal subgroup of finite index, K is a topological field, and the representation ρ is continuous in the sense that, in some (or, equivalently, in every) K–basis of the representation space, say V, of ρ, each matrix coefficient of ρ is a continuous K–valued function on G. Then each representation Λ_i of H is continuous. Indeed, in a suitable basis of V, $\rho|H$ is block diagonal, with blocks the Λ_i. So in this basis each matrix coefficient of each Λ_i is the restriction to H of a matrix coefficient of ρ, hence is continuous.

Chapter 4: Middle Convolution

4.0 Review of middle additive convolution: the class $\mathcal{P}_{\mathrm{conv}}$

(4.0.1) We fix a prime number ℓ. We work on $\mathbb{A}^1$ over an algebraically closed field k in which ℓ is invertible. We wish to define a certain class $\mathcal{P}_{\mathrm{conv}}$ of irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaves $\mathcal{F}$ on $\mathbb{A}^1$. Given an irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on $\mathbb{A}^1$ (or, equivalently, a nonpunctual irreducible perverse sheaf $K = \mathcal{F}[1]$ on $\mathbb{A}^1$), denote by

(4.0.1.1) $$S := \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$$

the finite set of points in $\mathbb{A}^1$ at which $\mathcal{F}$ is not lisse.

(4.0.2) We say that $\mathcal{F}$ lies in $\mathcal{P}_{\mathrm{conv}}$ if

(4.0.2.1) $$\mathrm{rank}(\mathcal{F}) + \#S + \Sigma_{t \text{ in } S\cup\{\infty\}} \mathrm{Swan}_t(\mathcal{F}) \geq 3.$$

(4.0.3) If k has characteristic zero, then among all irreducible middle extensions $\mathcal{F}$, only the constant sheaf $\overline{\mathbb{Q}}_\ell$ and the Kummer sheaves $\mathcal{L}_{\chi(x-\alpha)}$, ($\chi$ a nontrivial character of $\pi_1^{\mathrm{tame}}(\mathbb{G}_m)$, α in $\mathbb{A}^1(k)$) fail to lie in $\mathcal{P}_{\mathrm{conv}}$. Equivalently, in characteristic zero, an irreducible middle extension $\mathcal{F}$ lies in $\mathcal{P}_{\mathrm{conv}}$ if and only if $\#S \geq 2$.

(4.0.4) If k has characteristic $p > 0$, then among all irreducible middle extensions $\mathcal{F}$, only the constant sheaf $\overline{\mathbb{Q}}_\ell$, the Kummer sheaves $\mathcal{L}_{\chi(x-\alpha)}$ as above, and the Artin–Schreier sheaves $\mathcal{L}_{\psi(\alpha x)}$ (ψ a nontrivial additive character of $\mathbb{F}_p$, α in $\mathbb{A}^1(k)$) fail to lie in $\mathcal{P}_{\mathrm{conv}}$.

(4.0.5) In [Ka–RLS, 3.3.3 and 4.3.10, where the objects in $\mathcal{P}_{\mathrm{conv}}$ are called "of type 2d)"], it is shown that the class $\mathcal{P}_{\mathrm{conv}}$ is stable by middle additive convolution with Kummer sheaves $j_*\mathcal{L}_{\chi(x)}$ on $\mathbb{A}^1$, χ any nontrivial character of $\pi_1^{\mathrm{tame}}(\mathbb{G}_m)$, and j the inclusion of $\mathbb{G}_m$ into $\mathbb{A}^1$. Let us recall the basic setup. Given $\mathcal{F}$ in $\mathcal{P}_{\mathrm{conv}}$, and a Kummer sheaf $\mathcal{L}_{\chi(x)}$ as above, form the perverse sheaves $K := \mathcal{F}[1]$ and $L := j_*\mathcal{L}_{\chi(x)}[1]$ on $\mathbb{A}^1$. On $\mathbb{A}^2$ with its two projections to $\mathbb{A}^1$, form the external tensor product $K\boxtimes L := (\mathrm{pr}_1{}^*K)\otimes(\mathrm{pr}_2{}^*L)$. By the sum map

(4.0.5.1) $$\mathrm{sum}: \mathbb{A}^2 \to \mathbb{A}^1,$$

form the two flavors of total direct image, $R\mathrm{sum}_!(K\boxtimes L)$ and $R\mathrm{sum}_*(K\boxtimes L)$. Because $\mathcal{F}$ is in $\mathcal{P}_{\mathrm{conv}}$, both $R\mathrm{sum}_!(K\boxtimes L)$ and $R\mathrm{sum}_*(K\boxtimes L)$ are perverse. The middle additive convolution $K*_{\mathrm{mid}+}L$ is defined to be the image, in the category of perverse sheaves, of the canonical "forget supports" map:

(4.0.5.2) $$K*_{\mathrm{mid}+}L := \mathrm{Image}(R\mathrm{sum}_!(K\boxtimes L) \to R\mathrm{sum}_*(K\boxtimes L)).$$

One knows that $K*_{\mathrm{mid}+}L$ is of the form $\mathcal{G}[1]$ for an irreducible middle extension $\mathcal{G}$ in $\mathcal{P}_{\mathrm{conv}}$. We write

(4.0.5.3) $\mathcal{G} = \mathcal{F}*_{\mathrm{mid}+}\mathcal{L}_\chi$.

4.1 Effect on local monodromy

(4.1.1) We now recall the relations between the local monodromies of $\mathcal{F}$ and $\mathcal{G}$. For any point t in $\mathbb{P}^1$, we denote by $\mathcal{F}$(t) and $\mathcal{G}$(t) the representations of the inertia group I(t) attached to $\mathcal{F}$ and $\mathcal{G}$ respectively. Given any $\overline{\mathbb{Q}}_\ell$–representation M(t) of I(t), we have its direct sum break ("upper numbering") decomposition [Ka–GKM, 1.1] into I(t)–stable pieces

(4.1.1.1) $$\mathrm{M(t)} = \oplus_{\alpha \geq 0 \text{ in } \mathbb{Q}} \mathrm{M(t)(break}=\alpha).$$

If we collect the terms according as to whether α=0 or α >0, we get the coarser decomposition

(4.1.1.2) $$\mathrm{M(t)} = \mathrm{M(t)}^{\mathrm{tame}} \oplus \mathrm{M(t)}^{\mathrm{wild}}.$$

(4.1.2) Denote by Rep(I(t), $\overline{\mathbb{Q}}_\ell$) the category of finite–dimensional continuous $\overline{\mathbb{Q}}_\ell$–representations of I(t). For any subset $\mathcal{B}$ of $\mathbb{Q}_{\geq 0}$, denote by Rep(I(t), $\overline{\mathbb{Q}}_\ell$)(breaks in $\mathcal{B}$) the full subcategory of objects all of whose breaks lie in $\mathcal{B}$. When k has characteristic p > 0, Laumon [Lau–TFC, 2.4] has defined local Fourier transform functors

(4.1.2.1) $$\mathrm{FTloc(t, \infty) : Rep(I(t)}, \overline{\mathbb{Q}}_\ell) \to \mathrm{Rep(I(\infty)}, \overline{\mathbb{Q}}_\ell)$$

with the following properties.

(4.1.3) For t in $\mathbb{A}^1$, FTloc(t, ∞) is an equivalence

(4.1.3.1) $$\mathrm{FTloc(t, \infty) : Rep(I(t)}, \overline{\mathbb{Q}}_\ell) \cong \mathrm{Rep(I(\infty)}, \overline{\mathbb{Q}}_\ell)(\mathrm{breaks} \leq 1),$$

which interchanges objects of dimension b having all breaks a/b with objects of dimension a+b having all breaks a/(a+b).

(4.1.4) For t = ∞, FTloc(∞, ∞) kills Rep(I(∞), $\overline{\mathbb{Q}}_\ell$)(breaks ≤ 1), and induces an autoequivalence of Rep(I(∞), $\overline{\mathbb{Q}}_\ell$)(breaks > 1), which interchanges objects of dimension a having all breaks (a+b)/a wth objects of dimension b having all breaks (a+b)/b.

(4.1.5) In terms of these local Fourier transform functors, we can define, in characteristic p > 0, local convolution functors as follows.

(4.1.6) For t in $\mathbb{A}^1$, we define

(4.1.6.1) $$\mathrm{MC}_\chi\mathrm{loc(t) : Rep(I(t)}, \overline{\mathbb{Q}}_\ell) \to \mathrm{Rep(I(t)}, \overline{\mathbb{Q}}_\ell)$$

to be the autoequivalence

(4.1.6.2) $$\mathrm{FTloc(t, \infty)}^{-1}\circ(\mathrm{M} \mapsto \mathrm{M}\otimes\mathcal{L}_{\overline{\chi}(\mathrm{x})})\circ\mathrm{FTloc(t, \infty)},$$

where $\overline{\chi}$ denotes the inverse character. The local convolution functor $\mathrm{MC}_{\overline{\chi}}$loc(t) is a quasi–inverse to MC_χloc(t).

(4.1.7) For t=∞, we define

(4.1.7.1) $$\mathrm{MC}_\chi\mathrm{loc}(\infty) : \mathrm{Rep(I(\infty)}, \overline{\mathbb{Q}}_\ell)(\mathrm{breaks} > 1) \to \mathrm{Rep(I(\infty)}, \overline{\mathbb{Q}}_\ell)(\mathrm{breaks} > 1)$$

to be the autoequivalence

(4.1.7.2) $$\mathrm{FTloc}(\infty, \infty)^{-1}\circ(\mathrm{M} \mapsto \mathrm{M}\otimes\mathcal{L}_{\overline{\chi}(\mathrm{x})})\circ\mathrm{FTloc}(\infty, \infty).$$

Its quasi–inverse is $\mathrm{MC}_{\overline{\chi}}$loc(∞).

(4.1.8) These functors preserve both dimensions and breaks. On **tame** objects M in $\mathrm{Rep}(\mathrm{I}(t), \overline{\mathbb{Q}}_\ell)$, t in $\mathbb{A}^1$, $\mathrm{MC}_\chi\mathrm{loc}(t)$ is just the functor

$$M \mapsto M\otimes\mathcal{L}_{\chi(x-t)},$$

cf. [Ka–RLS, proof of 3.3.6] On objects which are not tame, $\mathrm{MC}_\chi\mathrm{loc}(t)$ is not given by this rule in general, cf. [Ka–RLS, 3.4] for a discussion of this point.

(4.1.9) In characteristic zero, we **define**, for t in $\mathbb{A}^1$, $\mathrm{MC}_\chi\mathrm{loc}(t)$ to be the functor on $\mathrm{Rep}(\mathrm{I}(t), \overline{\mathbb{Q}}_\ell)$ given by

$$M \mapsto M\otimes\mathcal{L}_{\chi(x-t)},$$

Using the relation of middle additive convolution to Fourier transform, Laumon's results on the local structure of Fourier transform, and, if the characteristic is zero, a "reduction to characteristic p" argument, we find

Theorem 4.1.10 [Ka–RLS, 3.3.5–6 and 4.3.11] Given $\mathcal{F}$ in $\mathcal{P}_{\mathrm{conv}}$ and a nontrivial Kummer sheaf $\mathcal{L}_{\chi(x)}$, put $\mathcal{G} := \mathcal{F}*_{\mathrm{mid}+}\mathcal{L}_\chi$ in $\mathcal{P}_{\mathrm{conv}}$.

1) For t in $\mathbb{A}^1$, the I(t)–representations $\mathcal{F}(t)$ and $\mathcal{G}(t)$ are related as follows:

$$\mathcal{G}(t)/\mathcal{G}(t)^{\mathrm{I}(t)} \cong \mathrm{MC}_\chi\mathrm{loc}(t)(\mathcal{F}(t)/\mathcal{F}(t)^{\mathrm{I}(t)}).$$

1a) We have an isomorphism of tame I(t)–representations

$$\mathcal{G}(t)^{\mathrm{tame}}/\mathcal{G}(t)^{\mathrm{I}(t)} \cong (\mathcal{F}(t)^{\mathrm{tame}}/\mathcal{F}(t)^{\mathrm{I}(t)})\otimes\mathcal{L}_{\chi(x-t)}.$$

1b) We have an equality of dimensions

$$\dim\mathcal{G}(t)^{\mathrm{wild}} = \dim\mathcal{F}(t)^{\mathrm{wild}}.$$

1c) We have an equality of dimensions

$$\dim \mathcal{G}(t)/\mathcal{G}(t)^{\mathrm{I}(t)} = \dim \mathcal{F}(t)/\mathcal{F}(t)^{\mathrm{I}(t)}.$$

2) The I(∞)–representations $\mathcal{F}(\infty)$ and $\mathcal{G}(\infty)$ are related as follows.

2a) There exists a tame I(∞)–representation M such that

$$\mathcal{F}(\infty)^{\mathrm{tame}} = M/M^{\mathrm{I}(\infty)},$$
$$\mathcal{G}(\infty)^{\mathrm{tame}} = (M\otimes\mathcal{L}_{\chi(x)})/(M\otimes\mathcal{L}_{\chi(x)})^{\mathrm{I}(\infty)}.$$

2b) We have an isomorphism of I(∞)–representations

$$\mathcal{G}(\infty)(0 < \mathrm{break} \le 1) \cong \mathcal{F}(\infty)(0 < \mathrm{break} \le 1).$$

2c) We have an isomorphism of I(∞)–representations

$$\dim \mathcal{G}(\infty)(\mathrm{breaks} > 1) = \mathrm{MC}_\chi\mathrm{loc}(\infty)(\mathcal{F}(\infty)(\mathrm{breaks} > 1)).$$

2d) We have an equality of dimensions

$$\dim \mathcal{G}(\infty)^{\mathrm{wild}} = \dim \mathcal{F}(\infty)^{\mathrm{wild}}.$$

proof If k has characteristic zero, then $\mathcal{F}$ and $\mathcal{G}$ are necessarily tame, and this is [Ka–RLS, 4.3.11], proven by reducing to the characteristic p > 0 case. If k has characteristic p > 0, this is just a spelling out of [Ka–RLS, 3.3.5], using the discussion in the proof of [Ka–RLS, 3.3.6] to identify more precisely what happens on the tame parts. QED

Corollary 4.1.11 Hypotheses and notations as in 4.1.10, the action of $I(\infty)$ on $\mathcal{G}(\infty)$ is **not** semisimple, and hence does **not** factor through a finite quotient of $I(\alpha)$, if any of the following conditions is satisfied.

1) $\mathcal{F}(\infty)^{I(\infty)} \neq 0$, i.e., $\mathcal{F}(\infty)^{\mathrm{tame}}$ has a unipotent Jordan block of dimension ≥ 1.

2) $\mathcal{F}(\infty)^{\mathrm{tame}} \otimes \mathcal{L}_{\overline{\rho}(x)}$ has a unipotent Jordan block of dimension ≥ 2, for some $\rho \neq \overline{\chi}$, ρ nontrivial.

3) $\mathcal{F}(\infty)^{\mathrm{tame}} \otimes \mathcal{L}_{\chi(x)}$ has a unipotent Jordan block of dimension ≥ 3.

4) $\mathcal{F}(\infty)^{\mathrm{wild}}$ is not $I(\infty)$–semisimple.

proof If 1) holds, then from the isomorphism $\mathcal{F}(\infty)^{\mathrm{tame}} = M/M^{I(\infty)}$ we see that M has a direct summand which is a unipotent Jordan block U of dimension ≥ 2. Then $U \otimes \mathcal{L}_{\chi(x)}$ is a direct summand of $M \otimes \mathcal{L}_{\chi(x)}$. But $(U \otimes \mathcal{L}_{\chi(x)})^{I(\infty)} = 0$, so $U \otimes \mathcal{L}_{\chi(x)}$ is a direct summand of $(M \otimes \mathcal{L}_{\chi(x)})/(M \otimes \mathcal{L}_{\chi(x)})^{I(\infty)} \cong$ of $\mathcal{G}(\infty)$.

If 2) holds, then M has a direct summand $U \otimes \mathcal{L}_{\rho(x)}$ with U a unipotent Jordan block of dimension ≥ 2, and hence $\mathcal{G}(\infty)$ has a direct summand $U \otimes \mathcal{L}_{\rho(x)} \otimes \mathcal{L}_{\chi(x)}$.

If 3) holds, then M has a direct summand $U \otimes \mathcal{L}_{\overline{\chi}(x)}$ with U a unipotent Jordan block of dimension $d \geq 3$. Hence $M \otimes \mathcal{L}_{\chi(x)}$ has a direct summand U, and hence $\mathcal{G}(\infty)^{\mathrm{tame}}$ has a direct summand $U/U^{I(\infty)}$, which is a unipotent Jordan block of dimension $d-1 \geq 2$.

Suppose 4) holds. If $\mathcal{F}(\infty)(0 < \text{slopes} \leq 1)$ is not $I(\infty)$–semisimple, neither is the isomorphic representation $\mathcal{G}(\infty)(0 < \text{slopes} \leq 1)$. Suppose $\mathcal{F}(\infty)(\text{slopes} > 1)$ is not $I(\infty)$–semisimple. As $MC_\chi loc(\infty)$ is an autoequivalence, it preserves non–semisimplicity, so $\mathcal{G}(\infty)(\text{slopes} > 1)$ is not $I(\infty)$–semisimple. QED

4.2 Calculation of $MC_\chi loc(\alpha)$ on certain wild characters

Proposition 4.2.1 Let k be an algebraically closed field of characteristic $p > 0$, α in $\mathbb{A}^1(k)$, ℓ a prime $\neq p$. Let χ and ρ be $(\overline{\mathbb{Q}}_\ell)^\times$–valued characters of $I(\alpha)$. Suppose that χ is nontrivial of order prime to p, and suppose that ρ is nontrivial of p–power order. Put $n := \mathrm{Swan}(\rho)$. Then for some nontrivial character $\tilde{\rho}$ of $I(\alpha)$ of p–power order and the same Swan conductor n, we have

$$MC_\chi loc(\alpha)(\rho) = \chi^{n+1}\tilde{\rho}.$$

proof By additive translation, we first reduce to the case $\alpha = 0$. We then use a global argument. Any character ρ of p–power order of $I(0)$ has a canonical extension to a character of p–power order of $\pi_1(\mathbb{P}^1 - (0)$, cf. [Ka–LG, 1.4.2]. View this canonical extension as a lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf on $\mathbb{P}^1 - \{0\}$, restrict it to $\mathbb{G}_m$, and denote by $\mathcal{F}$ in $\mathcal{P}_{\mathrm{conv}}$ its middle extension to $\mathbb{A}^1$. Denote by $\mathcal{H}$ in $\mathcal{P}_{\mathrm{conv}}$ the middle additive convolution

$$\mathcal{H} := \mathcal{F}*_{\mathrm{mid}+}\mathcal{L}_\chi.$$

Directly from the definitions, one sees that $\mathcal{H}$ is lisse on $\mathbb{G}_m$ of rank n+1.

Now apply the results on local monodromy of middle additive convolutions recalled in Theorem 4.1.10 above. We have an isomorphism of I(0)–representations

$$\mathcal{H}(0)/\mathcal{H}(0)^{\mathrm{I}(0)} = \mathrm{MC}_\chi\mathrm{loc}(0)(\rho).$$

Because $\mathrm{MC}_\chi\mathrm{loc}(0)$ preserves both dimensions and breaks, we see that $\mathcal{H}(0)/\mathcal{H}(0)^{\mathrm{I}}$ is a (one–dimensional) character of I(0), whose Swan conductor is n.

The local monodromy of $\mathcal{H}$ at ∞ is

$$\mathcal{H}(\infty) \cong \mathcal{L}_\chi\otimes(\text{unipotent pseudoreflection of size n+1}).$$

Now consider the lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\det(\mathcal{H})$ on $\mathbb{G}_m$. As I(0)–representation, it is $\mathrm{MC}_\chi\mathrm{loc}(0)(\rho) = \mathcal{H}(0)/\mathcal{H}(0)^{\mathrm{I}(0)}$ [simply because $\mathcal{H}(0)^{\mathrm{I}(0)}$ has codimension 1 in $\mathcal{H}(0)$]. As $\mathrm{I}(\infty)$–representation, it is $\mathcal{L}_{\chi^{n+1}}$. Hence $\mathcal{L}_{\chi^{-n-1}}\otimes\det(\mathcal{H})$ is lisse of rank one on $\mathbb{P}^1 - (0)$, so must have p–power order (because $\mathbb{P}^1 - (0)$ is tamely simply connected). Its restriction to I(0) is the required character $\tilde{\rho}$. QED

Corollary 4.2.2 Let k be an algebraically closed field of characteristic 2, α in $\mathbb{A}^1(\mathrm{k})$. Let χ and ρ be $(\overline{\mathbb{Q}}_\ell)^\times$–valued characters of $\mathrm{I}(\alpha)$. Suppose that χ is nontrivial of odd order, and suppose that ρ has order 2 and $\mathrm{Swan}(\rho) = 1$.Then for some nontrivial character $\tilde{\rho}$ of $\mathrm{I}(\alpha)$ of order 2 and Swan conductor 1, we have

$$\mathrm{MC}_\chi\mathrm{loc}(\alpha)(\rho) = \chi^2\tilde{\rho}.$$

Thus $\mathrm{MC}_\chi\mathrm{loc}(\alpha)(\rho)$ is a character of order $2\times(\text{order of }\chi) \geq 6$.

proof The only point to remark is that, in any finite characteristic p, nontrivial characters of $\mathrm{I}(\alpha)$ of p–power order having Swan conductor < p are all of order p, as one sees from [Ka–GKM, 8.5.7.1] and an obvious induction. Therefore $\tilde{\rho}$ has order 2. Since χ has odd order, χ^2 has the same odd order, whence the asserted order of $\mathrm{MC}_\chi\mathrm{loc}(\alpha)(\rho)$. QED

Chapter 5: Twist Sheaves and Their Monodromy

5.0 Families of twists: basic definitions and constructions

(5.0.1) In this section, we make explicit the "families of twists" we will be concerned with. We fix an algebraically closed field k, a proper smooth connected curve C/k whose genus is denoted g, and a prime number ℓ invertible in k. We also fix an integer $r \geq 1$, and an irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on C of generic rank r. This means that for some dense open set U in C, with $j : U \to C$ the inclusion, $\mathcal{F}|U$ is a lisse sheaf of rank r on U which is irreducible in the sense that the corresponding r–dimensional $\overline{\mathbb{Q}}_\ell$–representation of $\pi_1(U)$ is irreducible, and $\mathcal{F}$ on C is obtained from the lisse irreducible sheaf $\mathcal{F}|U$ on U by direct image: $\mathcal{F} \cong j_*(\mathcal{F}|U) := j_*j^*\mathcal{F}$.

(5.0.2) We say that $\mathcal{F}$ is self–dual if for every dense open set U on which it is lisse, $\mathcal{F}|U$ is self–dual as lisse sheaf, i.e., isomorphic to its contragredient. It is equivalent to say that the perverse sheaf $\mathcal{F}[1]$ on C is self–dual, but we will not need this more sophisticated point of view.

(5.0.3) The finite set of points of C at which $\mathcal{F}$ fails to be lisse, i.e., the set of points x for which the inertia group I(x) acts nontrivially on $\mathcal{F}$, will be denoted Sing($\mathcal{F}$), the set of "singularities" of $\mathcal{F}$. Thus $\mathcal{F}$ is lisse on C – Sing($\mathcal{F}$), and Sing($\mathcal{F}$) is minimal with this property.

(5.0.4) We fix an effective divisor $D = \sum a_i P_i$ on C, whose degree $d := \sum a_i$ satisfies $d \geq 2g+1$. Some or all or none of the points P_i may lie in Sing($\mathcal{F}$). We denote by L(D) the Riemann–Roch space $H^0(C, I^{-1}(D))$, and we view L(D) as a space of functions (maps to $\mathbb{A}^1$) on the open curve C – D.

(5.0.5) Corresponding to the choice of D as the "points at ∞" of C, we break up the set Sing($\mathcal{F}$) as the disjoint union

(5.0.5.1) $$\mathrm{Sing}(\mathcal{F}) := \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}} \amalg \mathrm{Sing}(\mathcal{F})_\infty$$

where

(5.0.5.2) $$\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}} := \mathrm{Sing}(\mathcal{F}) \cap (C-D),$$

(5.0.5.3) $$\mathrm{Sing}(\mathcal{F})_\infty := \mathrm{Sing}(\mathcal{F}) \cap D.$$

Lemma 5.0.6 Given a finite subset S of C–D, denote by

$$\mathrm{Fct}(C, d, D, S) \subset L(D)$$

the set of nonzero functions f in L(D) with the following property:

the divisor of zeroes of f, $f^{-1}(0)$, consists of d = degree(D) distinct points, none of which lies in S∪D. Then Fct(C, d, D, S) is (the set of k–points of) a dense open set *Fct*(C, d, D, S) in L(D) (viewed as the set of k–points of an affine space $\mathbb{A}^{d+1-g}$ over k).

proof The projective space $\mathbb{P}(L(D)^\vee)$ of lines in L(D) is the space of effective divisors of degree d which are linearly equivalent to D. In the space $\mathrm{Sym}^d(C)$ of all effective divisors of degree D, those consisting of d distinct points, none of which lies in S∪D, form an open set, say U_1. When we map $\mathrm{Sym}^d(C)$ to $\mathrm{Jac}^d(C)$, the fibre over the class of D is $\mathbb{P}(L(D)^\vee)$. The intersection of this fibre with U_1 is an open set U_2 in $\mathbb{P}(L(D)^\vee)$. The inverse image U_3 of this set in $L(D) - \{0\}$ is the set Fct(C, d, D, S) in L(D), which is thus open.

To see that U_3 is nonempty, we argue as follows. Suppose there exists a function f in L(D) whose divisor of poles is D and whose differential df is nonzero. Then for any t in k which is not a value taken by f on either S or on the set of zeroes in C–D of df, the function f–t lies in U_3 (it is nonzero on S, and it has simple zeroes because it has no zeroes in common with df).

Why does such an f exist? By Riemann–Roch, for each point P_i in D, $L(D - P_i)$ is a hyperplane in L(D): as k is infinite, L(D) is not the union of finitely many hyperplanes. So we can find a function f in L(D) whose divisor of poles is D. If any of the coefficients a_i in $D = \sum a_i P_i$ is invertible in k, then df is nonzero, because at P_i it has a pole of order $1+a_i$. If all a_i vanish in k, then k has charactertistic p, all the a_i are divisible by p, say $a_i = pb_i$, and $D = pD_0$, for D_0 the divisor $D_0 := \sum_i b_i P_i$. If df vanishes, then $f = g^p$ for some g in $L(D_0)$. In this case, pick a function g in $L(D - P_1)$ whose divisor of poles is $D - P_1$ (still possible by Riemann–Roch). Then dg is nonzero (it has a pole of order a_1 at P_1). For all but finite many values of t in k, f – tg still has divisor of poles D. For any such t, f – tg is the desired function. QED

Remark 5.0.7 Perhaps the simplest example to keep in mind is this. Take C to be $\mathbb{P}^1$, and take D to be $d\infty$. So here C–D is $\mathbb{A}^1 = \mathrm{Spec}(k[X])$, and *Fct*(C, d, D, S) is all the polynomials of degree d in one variable X with d distinct zeroes, none of which lies in S.

(5.0.8) We now turn to our final piece of data, a nontrivial $\overline{\mathbb{Q}}_\ell^\times$–valued character χ of finite order n ≥ 2 of the tame fundamental group of $\mathbb{G}_m/k$, corresponding to a lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}_\chi$ on $\mathbb{G}_m$. The order n of χ is necessarily invertible in k, indeed $\pi_1^{\mathrm{tame}}(\mathbb{G}_m/k)$ is the inverse limit of the groups $\mu_N(k)$ over those N invertible in k, corresponding to the various Kummer coverings $x \mapsto x^n$ of $\mathbb{G}_m$ by itself.

(5.0.9) When k has positive characteristic, the $\mathcal{L}_\chi$'s having given order n are obtained concretely as follows. Take any finite subfield $\mathbb{F}_q$ of k which contains the n'th roots of unity (i.e., $q\equiv 1 \bmod n$), and take a character $\chi : (\mathbb{F}_q)^\times \to \overline{\mathbb{Q}}_\ell^\times$ of order n. View $\mathbb{G}_m/\mathbb{F}_q$ as an $(\mathbb{F}_q)^\times$–torsor over itself by the map ("Lang isogeny")

(5.0.9.1) $$1 - \mathrm{Frob}_q : x \mapsto x^{1-q},$$

and push out this torsor by the character $\chi : (\mathbb{F}_q)^\times \to \overline{\mathbb{Q}}_\ell^\times$ to obtain a lisse rank one $\mathcal{L}_\chi$ on $\mathbb{G}_m/\mathbb{F}_q$. Its pullback to $\mathbb{G}_m/k$ is an $\mathcal{L}_\chi$ of the same order n on $\mathbb{G}_m/k$, and every $\mathcal{L}_\chi$ of order n on $\mathbb{G}_m/k$ is obtained this way.

(5.0.10) Given f in Fct(C, d, D, $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$), we may view f as mapping the open curve $C - D - f^{-1}(0)$ to $\mathbb{G}_m$, and we form the lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}_{\chi(f)} := f^*\mathcal{L}_\chi$ on $C - D - f^{-1}(0)$. When no ambiguity is likely, we will also denote by $\mathcal{L}_{\chi(f)}$ the extension by direct image of this sheaf to all of C. We then "twist" $\mathcal{F}$ by $\mathcal{L}_{\chi(f)}$. This means that we pass to the open set

$$j : C - D - f^{-1}(0) - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}} \subset C,$$

on which both $\mathcal{F}$ and $\mathcal{L}_{\chi(f)}$ are lisse, on that open set we form $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$, and then we take the direct image $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ to C. Notice that this twisted sheaf $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on C is itself an irreducible middle extension.

(5.0.11) Since at each point of $f^{-1}(0)$ and at each point of $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ one of the factors $\mathcal{F}$ or $\mathcal{L}_{\chi(f)}$ is lisse, the sheaf $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) \mid C-D$ is the literal tensor product $\mathcal{F}\otimes\mathcal{L}_{\chi(f)} \mid C-D$. Thus if we denote by $j_\infty: C - D \to C$ the inclusion, $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ as defined above is obtained from the literal tensor product $\mathcal{F}\otimes\mathcal{L}_{\chi(f)} \mid C-D$ by taking direct image across D^{red}:

$$j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) = j_{\infty *}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}).$$

This alternate interpretation will be used later, in 5.2.4 and 5.2.5.

(5.0.12) We then form the cohomology groups $H^i(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$ with coefficients in the twist $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$. Our eventual goal is to study the variation of these cohomology groups as f varies. But first we must establish some basic properties of these groups for a fixed f.

5.1 Basic facts about the groups $H^i(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$

Lemma 5.1.1 Hypotheses and notations as in 5.0.1, 5.0.4, 5.0.8, and 5.0.10 above, the cohomology groups $H^i(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$ vanish for $i\neq 1$.

proof The H^i vanish for cohomological dimension reasons for i not in [0, 2]. For i=0, we have

$$H^0(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})) := H^0(C - D - f^{-1}(0) - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, \mathcal{F}\otimes\mathcal{L}_{\chi(f)}).$$

This group vanishes because $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$ is lisse on the open curve, it is irreducible ($\mathcal{F}$ is irreducible, and $\mathcal{L}_{\chi(f)}$ has rank one) and nontrivial (because $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$ is nontrivially ramified at each of the d points of $f^{-1}(0)$). So the H^0 is the invariants in a nontrivial irreducible representation, so vanishes. Similarly, the birational invariance of H^2_c gives

$$H^2(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})) := H^2_c(C - D - f^{-1}(0) - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, \mathcal{F}\otimes\mathcal{L}_{\chi(f)}),$$

which is the Tate–twisted coinvariants in the same representation, so also vanishes. QED

(5.1.2) We next compute the dimension of $H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$, for f in $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$.

Given a point x in C(k), and a lisse sheaf $\mathcal{H}$ on some dense open set of C, we denote by $\mathcal{H}(x)$ the representation of I(x) given by $\mathcal{H}$ (strictly speaking, given by the pullback of $\mathcal{H}$ to the spectrum of the x–adic completion of the function field of C), and by $\mathcal{H}(x)^{I(x)}$, or simply $\mathcal{H}^{I(x)}$, the invariants in this representation. We will write $\mathcal{H}/\mathcal{H}^{I(x)}$ for $\mathcal{H}(x)/\mathcal{H}(x)^{I(x)}$. We will write

(5.1.2.1) $$\mathrm{drop}_x(\mathcal{H}) := \mathrm{drop}_x(\mathcal{H}(x)) := \dim(\mathcal{H}/\mathcal{H}^{I(x)}).$$

For any of the P_i occurring in $D = \sum a_iP_i$, and any f with divisor of poles D, the $I(P_i)$–representation $(\mathcal{L}_{\chi(f)})(P_i)$ depends only on χ^{a_i}, as follows. Choose a uniformizing parameter at P_i, and use it to identify the complete local ring of C at P_i with the complete local ring $k[[1/X]]$ (sic) of $\mathbb{P}^1$ at ∞, and to identify the inertia group $I(P_i)$ with $I(\infty)$. Consider the lisse sheaf $\mathcal{L}_{\chi^{a_i}} := \mathcal{L}_{\chi^{a_i}(X)}$ on $\mathbb{G}_m$. Then $(\mathcal{L}_{\chi(f)})(P_i)$ as $I(P_i)$–representation is just $(\mathcal{L}_{\chi^{a_i}})(\infty)$ as $I(\infty)$–representation. When we want to indicate unambiguously that we are thinking of $(\mathcal{L}_{\chi^{a_i}})(\infty)$ as an $I(P_i)$–representation by some choice of uniformizer as above, we will denote it $(\mathcal{L}_{\chi^{a_i}})(\infty, P_i)$.

Lemma 5.1.3 Hypotheses and notations as in 5.1.1 above, for any f in $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$, we have the dimension formula

(5.1.3.1)
$$\begin{aligned} h^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})) &= (2g-2 + \deg(D))\mathrm{rank}(\mathcal{F}) \\ &+ \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}). \\ &+ \textstyle\sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{drop}_s(\mathcal{F}) \\ &+ \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{drop}_{P_i}(\mathcal{F}(P_i)\otimes(\mathcal{L}_{\chi^{a_i}})(\infty, P_i)), \end{aligned}$$

and the inequality

(5.1.3.2) $$h^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})) \geq (2g-2 + \deg(D))\mathrm{rank}(\mathcal{F}) + \#\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}.$$

proof The inequality 5.1.3.2 is an immediate consequence of the asserted dimension formula 5.1.3.1 and the observation that $\mathrm{drop}_s(\mathcal{F}) \geq 1$ at each point in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. By Lemma 5.1.1, we have

$$h^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})) = -\chi(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})).$$

At each of the deg(D) distinct zeroes of f, $\mathcal{F}$ is lisse and $\mathcal{L}_{\chi(f)}$ is ramified, so $-\chi(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$ is equal to

$$= -\chi_c(C - f^{-1}(0) - D - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, \mathcal{F}\otimes\mathcal{L}_{\chi(f)})$$
$$-\Sigma_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \dim(\mathcal{F}(s)^{I(s)})$$
$$- \Sigma_{P_i \text{ in } D^{\mathrm{red}}} \dim((\mathcal{F}(P_i)\otimes(\mathcal{L}_{\chi}a_i)(\infty, P_i))^{I(P_i)}).$$

Now use the Euler–Poincaré formula to write this as

$$= (2g-2 + \deg(D) + \#D^{\mathrm{red}} + \#\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})\mathrm{rank}(\mathcal{F})$$
$$+ \Sigma_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \Sigma_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}).$$
$$-\Sigma_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \dim(\mathcal{F}(s)^{I(s)})$$
$$- \Sigma_{P_i \text{ in } D^{\mathrm{red}}} \dim((\mathcal{F}(P_i)\otimes(\mathcal{L}_{\chi}a_i)(\infty, P_i))^{I(P_i)})$$
$$= (2g-2 + \deg(D))\mathrm{rank}(\mathcal{F})$$
$$+ \Sigma_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \Sigma_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}).$$
$$+ \Sigma_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{drop}_s(\mathcal{F})$$
$$+ \Sigma_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{drop}_{P_i}(\mathcal{F}(P_i)\otimes(\mathcal{L}_{\chi}a_i)(\infty, P_i)). \qquad \text{QED}$$

5.2 Putting together the groups $H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$

Construction–Proposition 5.2.1 (compare [Ka–RLS, 2.7.2]) Hypotheses and notations as in 5.0.1, 5.0.4, 5.0.8, and 5.0.10 above, There is a natural lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{G}$ on the space

$$Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$$

whose stalk at f is the cohomology group $H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$. More precisely, over the parameter space

$$X := Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}),$$

consider the proper smooth curve $\mathcal{C} := C\times X$, and in it the relative divisor $\mathcal{D}$ defined at "time f" by $D^{\mathrm{red}} + \mathrm{Sing}(\mathcal{F})\mathrm{finite} + f^{-1}(0)$. Then $\mathcal{D}$ is finite etale over the base of constant degree

$$\#(D^{\mathrm{red}}) + \#(\mathrm{Sing}(\mathcal{F})\mathrm{finite}) + d.$$

On $\mathcal{C} - \mathcal{D}$, we have the lisse sheaf $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$. Denote the projection

$$\pi : \mathcal{C} - \mathcal{D} \to Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}).$$

We have the following results.

1) The sheaves $R^i\pi_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ vanish for $i\neq 1$, and $R^1\pi_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is lisse.

2) The sheaves $R^i\pi_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ vanish for $i\neq 1$, and

$R^1\pi_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is lisse, and of formation compatible with arbitrary change of base.

3) The image $\mathcal{G}$ of the natural "forget supports" map

$$R^1\pi_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) \to R^1\pi_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$$

is lisse, of formation compatible with arbitrary change of base. The stalk of $\mathcal{G}$ at the k−valued point "f" of *Fct*(C, d, D, $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$) is the cohomology group $H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$.

4) If the irreducible middle extension $\mathcal{F}$ on C is orthogonally (respectively symplectically) self−dual, and χ has order two, then the lisse sheaf $\mathcal{G}$ on X is symplectically (respectively orthogonally) self−dual.

5) The rank of $\mathcal{G}$ is equal to

$$\begin{aligned}\mathrm{rank}(\mathcal{G}) = {} & (2g-2 + \deg(D))\mathrm{rank}(\mathcal{F}) \\ & + \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}) \\ & + \textstyle\sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{drop}_s(\mathcal{F}) \\ & + \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{drop}_{P_i}(\mathcal{F}(P_i)\otimes(\mathcal{L}_{\chi}a_i)(\infty, P_i)).\end{aligned}$$

6) We have the inequality

$$\mathrm{rank}(\mathcal{G}) \geq (2g-2 + \deg(D))\mathrm{rank}(\mathcal{F}) + \#\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}.$$

proof 1) By proper base change and the previous lemma, we have the vanishing of the $R^i\pi_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ for $i\neq 1$. To show that $R^1\pi_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is lisse, we apply Deligne's semicontinuity theorem [Lau−SC, 2.1.2], according to which it suffices to show the $\mathbb{Z}$−valued function which attaches to each k−valued point "f" of the base the sum of the Swan conductors of $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$ at all the points at infinity,

$$\begin{aligned}f \mapsto {} & \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) \\ & + \textstyle\sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) \\ & + \textstyle\sum_{x \text{ in } f^{-1}(0)} \mathrm{Swan}_x(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}),\end{aligned}$$

is constant. As $\mathcal{L}_{\chi(f)}$ is rank one and everywhere tame, and $\mathcal{F}$ is lisse at every point of $f^{-1}(0)$, the terms at points of $f^{-1}(0)$ all vanish, and those at other points don't see the $\mathcal{L}_{\chi(f)}$. Thus the function is equal to the constant

$$\textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}).$$

Assertion 2) results by Poincaré duality from 1) for the dual sheaf $\mathcal{F}^\vee\otimes\mathcal{L}_{\overline{\chi}(f)}$. Once we have 1) and 2), $\mathcal{G}$ is lisse and of formation compatible with arbitrary change of base, being the image of a map of such sheaves on a smooth base X. That $\mathcal{G}$ has the asserted stalk at "f" amounts, by base change, to the fact that $H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))$ is the image of the "forget supports" map

$$H^1_c(C - D - f^{-1}(0) - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, \mathcal{F}\otimes\mathcal{L}_{\chi(f)})$$
$$\to H^1(C - D - f^{-1}(0) - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, \mathcal{F}\otimes\mathcal{L}_{\chi(f)}).$$

Assertion 4) results from 1), 2), and 3), by Poincaré duality and standard properties of cup product. Because $\mathcal{G}$ is lisse, assertions 5) and 6) result from Lemma 5.1.3, applied to any single f in the parameter space $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$. QED

Notation 5.2.2 When we want to keep in mind the twist genesis of the lisse sheaf $\mathcal{G}$ on $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ constructed in 5.2.1 above, we will denote it $\mathrm{Twist}_{\chi,C,D}(\mathcal{F})$:

(5.2.2.1) $$\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F}).$$

Remark 5.2.3 It will also be convenient to have the following variant on the above description of the sheaf $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ on the space

$$X := Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}).$$

Start as before with the lisse irreducible sheaf $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$ on $\mathcal{C} - \mathcal{D}$. The base X is itself lisse, of dimension d + 1 − g, so $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]$ is perverse irreducible on $\mathcal{C}-\mathcal{D}$. Denote by

$$j : \mathcal{C} - \mathcal{D} \to \mathcal{C}$$

the inclusion, and form the middle extension $j_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g])$. Then according to [Ka–RLS, 2.7.2], if we denote by $\bar{\pi} : \mathcal{C} \to X$ the projection, we have

$$\mathcal{G}[d+1-g] = R\bar{\pi}_* j_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g])$$
$$= \mathrm{image}(R\pi_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]) \to R\pi_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g])),$$

where the image is taken in the category of perverse sheaves on X.

Lemma 5.2.4 With the notations of 5.2.1, denote by

$$j_1 : \mathcal{C} - \mathcal{D} \to \mathcal{C} - D^{\mathrm{red}}\times X = (C - D)\times X$$

the inclusion. Then the middle extension of $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]$ by j_1 is the [shifted] literal tensor product

$$(j_1)_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]) = \mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]$$

on (C−D)×X. Its formation commutes with arbitrary change of base on X.

proof We are forming the middle extension across two disjoint smooth divisors in (C − D)×X, namely f=0 and $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}\times X$. Consider the inclusions

$$j_2 : \mathcal{C} - \mathcal{D} \to \mathcal{C} - D^{\mathrm{red}}\times X - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}\times X,$$
$$j_3 : \mathcal{C} - D^{\mathrm{red}}\times X - \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}\times X \to \mathcal{C} - D^{\mathrm{red}}\times X.$$

Under j_2, we are extending across the divisor f=0. The sheaf $\mathcal{F}$ is lisse on the target

$(\mathcal{C} - \mathrm{D}^{\mathrm{red}}\times\mathrm{X}) - (\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}\times\mathrm{X})$, so we have

$$(j_2)_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]) \cong \mathcal{F}\otimes(j_2)_{!*}(\mathcal{L}_{\chi(f)}[d+2-g]).$$

To see that $(j_2)_{!*}(\mathcal{L}_{\chi(f)}[d+2-g]) = (j_{2!}\mathcal{L}_{\chi(f)})[d+2-g]$ amounts to showing that $j_{2*}\mathcal{L}_{\chi(f)}$ vanishes on f=0 (for then $j_{2*}\mathcal{L}_{\chi(f)}$ is lisse on f=0, and hence $(j_2)_{!*}(\mathcal{L}_{\chi(f)}[d+2-g]) = (j_{2*}\mathcal{L}_{\chi(f)})[d+2-g]$, but this latter is $(j_{2!}\mathcal{L}_{\chi(f)})[d+2-g]$). But near any point of f=0, f is part of a system of coordinates (f, coordinates for X), so by the Kunneth formula we are reduced to the fact that for $j : \mathbb{G}_m \to \mathbb{A}^1$ the inclusion, we have $j_!\mathcal{L}_\chi \cong j_*\mathcal{L}_\chi$.

When we extend by j_3, across $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}\times\mathrm{X}$, $\mathcal{L}_{\chi(f)}$ is lisse in a neighborhood of this divisor, we may pull it out, and then we are reduced, by Kunneth, to the fact that $\mathcal{F}$ on C−D is its own middle extension across $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. QED

Variant Construction of $\mathcal{G} := \mathrm{Twist}_{\chi,\mathrm{C},\mathrm{D}}(\mathcal{F})$ 5.2.5 (compare [Ka–RLS, 2.7.2]) Notations as in 5.2.1 above, over the parameter space

$$\mathrm{X} := \mathit{Fct}(\mathrm{C}, \mathrm{d}, \mathrm{D}, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}),$$

consider the proper smooth curve $\mathcal{C} := \mathrm{C}\times\mathrm{X}$ over X and in it the product divisor $\mathrm{D}^{\mathrm{red}}\times\mathrm{X}$. On the open curve $\mathcal{C} - \mathrm{D}^{\mathrm{red}}\times\mathrm{X} = (\mathrm{C} - \mathrm{D})\times\mathrm{X}$, form the literal tensor product sheaf $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$. Denote by

$$j_\infty : (\mathrm{C} - \mathrm{D})\times\mathrm{X} \to \mathrm{C}\times\mathrm{X}$$

the inclusion.
Denote by

$$\mathrm{pr}_2 : (\mathrm{C} - \mathrm{D})\times\mathrm{X} \to \mathrm{X} = \mathit{Fct}(\mathrm{C}, \mathrm{d}, \mathrm{D}, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$$

and

$$\bar{\pi} : \mathrm{C}\times\mathrm{X} \to \mathrm{X}$$

the projections. Then

1) The sheaves $\mathrm{R}^i\mathrm{pr}_{2!}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on $\mathit{Fct}(\mathrm{C}, \mathrm{d}, \mathrm{D}, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ vanish for $i\neq1$, and $\mathrm{R}^1\mathrm{pr}_{2!}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is lisse.

2) The sheaves $\mathrm{R}^i\mathrm{pr}_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on $\mathit{Fct}(\mathrm{C}, \mathrm{d}, \mathrm{D}, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ vanish for $i\neq1$, and $\mathrm{R}^1\mathrm{pr}_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is lisse, and of formation compatible with arbitrary change of base.

3) The perverse object $\mathcal{G}[d+1-g]$ on X is given by

$$\mathcal{G}[d+1-g] = \mathrm{R}\bar{\pi}_*j_{\infty!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g])$$
$$= \mathrm{image}(\mathrm{Rpr}_{2!}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})[d+2-g] \to \mathrm{Rpr}_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})[d+2-g]).$$

proof For 1), we see the vanishing fibre by fibre. The lisseness results from part 1) of the 5.2.1 via the long cohomology sequence for $\mathrm{Rpr}_{2!}$ attached to the short exact sequence of sheaves

$$0 \to j_{1!}j_1{}^*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) \to \mathcal{F}\otimes\mathcal{L}_{\chi(f)} \to \mathcal{F}\otimes\mathcal{L}_{\chi(f)} | (\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}\times X \to 0.$$

For 2), denote by $\mathcal{F}^\vee$ the middle extension sheaf dual to $\mathcal{F}$. By Lemma 5.2.4 above, applied to $\mathcal{F}^\vee$ and $\overline{\chi}$, $\mathcal{F}^\vee\otimes\mathcal{L}_{\overline{\chi}(f)}[d+2-g]$ is its own middle extension from $\mathcal{C}-\mathcal{D}$, so it is the Verdier dual of $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]$. So 2) for $\mathcal{F}\otimes\mathcal{L}_{\chi(f)}$ results from 1) for $\mathcal{F}^\vee\otimes\mathcal{L}_{\overline{\chi}(f)}$ by Poincaré duality. For 3), we already know (5.2.3) that

$$\mathcal{G}[d+1-g] = R\overline{\pi}_* j_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g])$$

for j the inclusion of $\mathcal{C} - \mathcal{D}$ into $\mathcal{C}$. So by the transitivity of middle extension ($j_{!*} = j_{\infty!*}\circ j_{1!*}$) and Lemma 5.2.4, we get

$$\mathcal{G}[d+1-g] = R\overline{\pi}_* j_{\infty!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g]).$$

That $R\overline{\pi}_* j_{\infty!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}[d+2-g])$ is the image of the canonical map

$$Rpr_{2!}((\mathcal{F}\otimes\mathcal{L}_{\chi(f)})[d+2-g]) \to Rpr_{2*}((\mathcal{F}\otimes\mathcal{L}_{\chi(f)})[d+2-g])$$

is [Ka–RLS, 2.7.2]. QED

5.3 First properties of twist families: relation to middle additive convolution on $\mathbb{A}^1$

(5.3.1) We begin with a direct image formula, which, although elementary, is a fundamental reduction tool in what is to follow.

(5.3.2) Fix f in $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$. Thus f is a finite flat map from C–D to $\mathbb{A}^1 = \mathrm{Spec}(k[X])$ of degree d, whose fibre over 0 consists of d distinct points, none of which lies in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. Denote by $\mathrm{CritPt}(f) \subset C-D$ the finite set of points in C–D at which df vanishes. Define

(5.3.2.1) $\qquad \mathrm{CritVal}(f, \mathcal{F}) := f(\mathrm{CritPt}(f)) \cup f(\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$,

a finite subset of $\mathbb{A}^1$. Then for t in $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$, the function t–f lies in $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$, and so we have a morphism

(5.3.2.2) $\qquad \mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F}) \to \mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$

given by $t \mapsto t-f$.

(5.3.3) What is the relation to convolution? We first explain the idea. For a good value t_0 of t, the stalk of $\mathcal{G}$ at t_0-f is the cohomology group

$$H^1(C, j_{\infty*}(\mathcal{F}\otimes\mathcal{L}_{\chi(t_0 - f)})) = \text{image of the "forget supports" map}$$
$$H_c{}^1(C-D, \mathcal{F}\otimes\mathcal{L}_{\chi(t_0 - f)}) \to H^1(C-D, \mathcal{F}\otimes\mathcal{L}_{\chi(t_0 - f)}).$$

Compute these cohomology groups on C–D by first mapping C–D to $\mathbb{A}^1$ by f. Since $\mathcal{L}_{\chi(t_0 - f)}$ is $f^*\mathcal{L}_{\chi(t_0-X)}$, the projection formula gives

$$H_c{}^1(C-D, \mathcal{F}\otimes\mathcal{L}_{\chi(t_0 - f)}) = H_c{}^1(\mathbb{A}^1, (f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t_0 - X)}),$$
$$H^1(C-D, \mathcal{F}\otimes\mathcal{L}_{\chi(t_0 - f)}) = H^1(\mathbb{A}^1, (f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t_0 - X)}).$$

So we get

$$H^1(C, j_{\infty*}(\mathcal{F}\otimes\mathcal{L}_{\chi(t_0 - f)}) = \text{image of the "forget supports" map}$$
$$H_c{}^1(\mathbb{A}^1, (f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t_0 - X)}) \to H^1(\mathbb{A}^1, (f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t_0 - X)}).$$

If we denote by $j_\infty : \mathbb{A}^1 \to \mathbb{P}^1$ the inclusion, this image is just $H^1(\mathbb{P}^1, j_{\infty*}((f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t_0 - X)}))$. According to [Ka–RLS, 2.8.5], there is an open dense set in $\mathbb{A}^1$ such that for t_0 in this open dense set, $H^1(\mathbb{P}^1, j_{\infty*}((f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t_0 - X)}))$ is the stalk at t_0 of the [shifted] middle additive convolution of $f_*\mathcal{F}$ with $\mathcal{L}_\chi$.

(5.3.4) Here is the precise result.

Proposition 5.3.5 Hypotheses and notations as in 5.2.1, fix f in Fct(C, d, D, $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$), viewed as a map from C–D to $\mathbb{A}^1$. Form the direct image sheaf $f_*(\mathcal{F}|C-D)$ on $\mathbb{A}^1$. The object

$$f_*(\mathcal{F}|C-D)[1]$$

on $\mathbb{A}^1$ is perverse. For $j : \mathbb{G}_m \to \mathbb{A}^1$ the inclusion, form the sheaf $j_*\mathcal{L}_\chi = j_!\mathcal{L}_\chi$ on $\mathbb{A}^1$, and the perverse object $j_*\mathcal{L}_\chi[1]$ on $\mathbb{A}^1$. Consider the middle additive convolution [Ka–RLS, 2.9]

$$f_*(\mathcal{F}|C-D)[1]*_{\mathrm{mid}+}j_*\mathcal{L}_\chi[1]$$

on $\mathbb{A}^1$. On $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$ we have a canonical isomorphism

$$([t \to t - f]^*\mathcal{G})[1] \cong (f_*(\mathcal{F}|C-D)[1])*_{\mathrm{mid}+}j_*\mathcal{L}_\chi[1].$$

proof The sheaf $\mathcal{F}$ on C–D is a middle extension, so $\mathcal{F}[1]$ on C–D is perverse. Since f is a finite map, f_*(perverse) is perverse.

We use the description of $\mathcal{G}[d+1-g]$ as

$$\mathrm{image}(Rpr_{2!}((\mathcal{F}\otimes\mathcal{L}_{\chi(f)})[d+1-g]) \to Rpr_{2*}((\mathcal{F}\otimes\mathcal{L}_{\chi(f)})[d+1-g]))$$

on *Fct*(C, d, D, $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$).

This description commutes with arbitrary change of base, so $([t \to t - f]^*\mathcal{G})[1]$ is

$$\mathrm{image}(Rpr_{2!}((\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)})[1]) \to Rpr_{2*}((\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)})[1])),$$

pr_2 the projection of $(C-D)\times(\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F}))$ to $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$. Now factor this projection the composition of

$$f\times id: (C-D)\times(\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})) \to \mathbb{A}^1\times(\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F}))$$

with the projection

$$\mathrm{pr}_{2,\mathbb{A}} : \mathbb{A}^1\times(\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F}))\to (\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})).$$

Since f is finite, we have $f_! = f_* = Rf_*$. The key point is that

$$\mathcal{L}_{\chi(t-f)}= (f\times id)^*\mathcal{L}_{\chi(t-X)}$$

and hence by the projection formula we find

$$\begin{aligned}Rpr_{2!}(\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)}) &= Rpr_{2!}(\mathcal{F}\otimes(f\times id)^*\mathcal{L}_{\chi(t-X)})\\ &= Rpr_{2,\mathbb{A}!}((f\times id)_!(\mathcal{F}\otimes(f\times id)^*\mathcal{L}_{\chi(t-X)}))\\ &= Rpr_{2,\mathbb{A}!}((f_!\mathcal{F})\otimes\mathcal{L}_{\chi(t-X)})\\ &= Rpr_{2,\mathbb{A}!}((f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t-X)}).\end{aligned}$$

Similarly we find

$$Rpr_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)}) = Rpr_{2,\mathbb{A}*}((f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t-X)}).$$

Thus we get that $([t \to t - f]^*\mathcal{G})[1]$ is

$$\begin{aligned}&\mathrm{image}(Rpr_{2!}((\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)})[1]) \to Rpr_{2*}((\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)})[1]))\\ &= \mathrm{image}(Rpr_{2,\mathbb{A}!}((f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t-X)})[1]) \to Rpr_{2,\mathbb{A}*}((f_*\mathcal{F})\otimes\mathcal{L}_{\chi(t-X)})[1])).\end{aligned}$$

This last image is the restriction to $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$ of the middle additive convolution of $f_*\mathcal{F}$ and $\mathcal{L}_\chi$, thanks to [Ka–RLS, 2.7.2 and 2.8.4]. QED

Proposition 5.3.6 Hypotheses and notations as in 5.2.1, suppose we are in one of the following situations:

1a) $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ is nonempty, $\deg(D) \ge 2g+1$, and $\mathrm{char}(k) \ne 2$.

1b) $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ is nonempty, $\deg(D) \ge 2g+3$, and $\mathrm{char}(k) = 2$.

2a) $\deg(D) \ge 4g+2$, and $\mathrm{char}(k) \ne 2$.

2b) $\deg(D) \ge 4g+6$, and $\mathrm{char}(k) = 2$.

Then the lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{G}$ on $\mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ is irreducible (or zero).

proof Suppose first that $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ is nonempty. If $\mathrm{char}(k) \ne 2$ [resp. if $\mathrm{char}(k) = 2$] pick a function f in $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ which also lies in the dense open set U of Theorem 2.2.6 [resp. Theorem 2.4.2], applied with S taken to be $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. Thus f as map from C–D to $\mathbb{A}^1$ is of Lefschetz type, and for each s in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$, the fibre $f^{-1}(s)$ consists of d distinct points, only one of which lies in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. By the Irreducible Induction Criterion 3.3.1, $f_*(\mathcal{F}|C-D)$ is an irreducible middle extension on $\mathbb{A}^1$. By [Ka–RLS, 2.9.7], the middle additive convolution $(f_*(\mathcal{F}|C-D)[1])*_{\mathrm{mid}+}j_*\mathcal{L}_\chi[1]$ on $\mathbb{A}^1$ is perverse irreducible. Hence its restriction to any dense open set of $\mathbb{A}^1$ is perverse irreducible (or zero).

We now turn to the case in which either $\mathrm{char}(k) \ne 2$ and $\deg(D) \ge 4g+2$, or $\mathrm{char}(k) = 2$ and

$\deg(D) \geq 4g+6$. Write D as the sum of two effective divisors $D = D_1 + D_2$, with both D_i having degree $\geq 2g+1$ (resp. $\geq 2g+3$ if $\mathrm{char}(k) = 2$).

Since $\deg(D_1) \geq 2g+1$ (resp. $\geq 2g+3$ if $\mathrm{char}(k) = 2$), we may choose a function f_1 in $\mathrm{Fct}(C, \deg(D_1), D_1, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}})$. Thus f_1 lies in $L(D_1)$, its divisor of poles is D_1, and it has $\deg(D_1)$ distinct zeroes, none of which lies in either $\mathrm{Sing}(\mathcal{F})$ or in D. Fix one such f_1.

As $\deg(D_2) \geq 2g+1$ if $\mathrm{char}(k) \neq 2$ [resp. $\geq 2g+3$ if $\mathrm{char}(k) = 2$], we may pick a function f_2 in $\mathrm{Fct}(C, \deg(D_2), D_2, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}}\cup f_1{}^{-1}(0))$ which lies in the open set U of Theorem 2.2.6 if $\mathrm{char}(k) \neq 2$ [resp. in the open set U of Theorem 2.4.2 if $\mathrm{char}(k) = 2$] with respect to S the set $f_1{}^{-1}(0)\cup(D^{\mathrm{red}} - D_2{}^{\mathrm{red}})$.

Thus f_2 has divisor of poles D_2, it has $\deg(D_2)$ distinct zeroes, none of which lies in $\mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}}\cup f_1{}^{-1}(0)$, and for each zero α of f_1, the f_2–fibre containing it, $f_2{}^{-1}(f_2(\alpha))$, consists of $\deg(D_2)$ distinct points, of which only α is a zero of f_1, and none of which lies in D. For any such f_2, the product $f_1 f_2$ lies in the space $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$. Moreover, for most scalars t, the product $f_1(t - f_2)$ lies in the space $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$. Thus for fixed f_1 and f_2, we have a map

$$\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}) \to \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}),$$
$$t \mapsto f_1(t - f_2).$$

Proposition 5.3.7 Given an effective D of degree $d \geq 4g+2$ (resp. $d \geq 4g+6$ if $\mathrm{char}(k) = 2$), write it as D_1+D_2 with both D_i effective of $\deg(D_i) \geq 2g+1$ (resp. $\geq 2g+3$ if $\mathrm{char}(k) = 2$). Fix

$$f_1 \text{ in } \mathrm{Fct}(C, \deg(D_1), D_1, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}}).$$

Fix a function f_2 in $\mathrm{Fct}(C, \deg(D_2), D_2, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}}\cup f_1{}^{-1}(0))$ which also lies in the open set U of Theorem 2.2.6 if $\mathrm{char}(k) \neq 2$ [resp. in the open set U of Theorem 2.4.2 if $\mathrm{char}(k) = 2$] with respect to the set $S := f_1{}^{-1}(0)\cup(\mathrm{Sing}(\mathcal{F})\cap(C-D_2))$. View f_2 as a finite flat map from $C - D_2$ to $\mathbb{A}^1$. For i=1,2, denote by

$$j_i : C - D \to C - D_i$$

the inclusion. Start with the sheaf $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$ on C– D, form its direct image $j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ on $C-D_2$, and take its direct image $f_{2*}j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ on $\mathbb{A}^1$. The object $f_{2*}j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]$ on $\mathbb{A}^1$ is perverse. For $j : \mathbb{G}_m \to \mathbb{A}^1$ the inclusion, form the sheaf $j_*\mathcal{L}_\chi = j_!\mathcal{L}_\chi$ on $\mathbb{A}^1$, and the perverse object $j_*\mathcal{L}_\chi[1]$ on $\mathbb{A}^1$. Consider the middle additive convolution [Ka–RLS, 2.9]

$$f_{2*}j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]*_{\mathrm{mid}+}j_*\mathcal{L}_\chi[1]$$

on $\mathbb{A}^1$. On $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$, we have a canonical isomorphism

$$([t \mapsto f_1(t - f_2)]^*\mathcal{G})[1] \cong (f_{2*}j_{2*}j_1{}^*(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1])*_{\mathrm{mid}+}j_*\mathcal{L}_\chi[1].$$

proof of 5.3.7 We work over the space

$$T := \mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}).$$

For i=1, 2, denote by $j_{i,\infty}$ the inclusion

$$j_{i,\infty}: C - D_i \to C.$$

We know that $([t \mapsto f_1(t - f_2)]^*\mathcal{G})[1]$ on T is given in terms of the projections

$$pr_{2,D} : (C-D)\times T \to T$$

and

$$pr_2 : C\times T \to T$$

as

$$\mathrm{image}(Rpr_{2,D!}((\mathcal{F}\otimes\mathcal{L}_{\chi(f_1(t-f_2))})[2]) \to Rpr_{2,D*}((\mathcal{F}\otimes\mathcal{L}_{\chi(f_1(t-f_2))})[2]))$$
$$= Rpr_{2*}((j_\infty\times id)_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1(t-f_2))}[2]))$$
$$= Rpr_{2*}((j_{2,\infty}\times id)_{!*}(j_2\times id)_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1(t-f_2))}[2])).$$

Now $(j_2\times id)_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1(t-f_2))}[1])$ means extending across points which are in D_1 but not in D_2, and $\mathcal{L}_{\chi(t-f_2)}$ is lisse near such points. So

$$(j_2\times id)_{!*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1(t-f_2))}[1]) = j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]\otimes\mathcal{L}_{\chi(t-f_2)}.$$

Thus $([t \mapsto f_1(t - f_2)]^*\mathcal{G})[1]$ on T is

$$Rpr_{2*}(j_{2,\infty}\times id)_{!*}((j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]\otimes\mathcal{L}_{\chi(t-f_2)}[1]).$$

Denote $\overline{f_2} := f_2$ viewed as a map of C to $\mathbb{P}^1$. Compute Rpr_{2*} by factoring pr_2 as

$$\overline{f_2}\times id: C\times T \to \mathbb{P}^1\times T$$

followed by

$$pr_{2,\mathbb{P}}: \mathbb{P}^1\times T \to T.$$

Thus

$$Rpr_{2*}(j_{2,\infty}\times id)_{!*}(j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]\otimes\mathcal{L}_{\chi(t-f_2)}[1]).$$
$$= Rpr_{2,\mathbb{P}*}(f_2\times id)_*(j_{2,\infty}\times id)_{!*}(j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]\otimes\mathcal{L}_{\chi(t-f_2)}[1])$$

In terms of the inclusion

$$j_{\mathbb{A}} : \mathbb{A}^1 \to \mathbb{P}^1$$

and

$$f_2 : C-D_2 \to \mathbb{A}^1$$

we have a cartesian diagram

$$\begin{array}{ccc} C - D_2 & \xrightarrow{j_{2,\infty}} & C \\ f_2 \downarrow & & \downarrow \overline{f_2} \\ \mathbb{A}^1 & \xrightarrow[j_{\mathbb{A}}]{} & \mathbb{P}^1 \end{array}$$

in which the horizontal maps are affine open immersions, and the vertical maps are finite. So we have

$$(\overline{f_2}\times id)_*(j_{2,\infty}\times id)_{!*} = (j_{\mathbb{A}}\times id)_{!*}(f_2\times id)_{!*}.$$

So we get

$$Rpr_{2,\mathbb{P}*}(\overline{f_2}\times id)_*(j_{2,\infty}\times id)_{!*}(j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]\otimes\mathcal{L}_{\chi(t-f_2)}[1])$$

$$= Rpr_{2,\mathbb{P}*}(j_{\mathbb{A}}\times id)_{!*}(f_2\times id)_*(j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1]\otimes\mathcal{L}_{\chi(t-f_2)}[1]).$$

Now $\mathcal{L}_{\chi(t-f_2)}$ is $f_2^{\ *}\mathcal{L}_{\chi(t-X)}$, so by the projection formula we may rewrite this last expression as

$$= Rpr_{2,\mathbb{P}*}(j_{\mathbb{A}}\times id)_{!*}(\mathcal{L}_{\chi(t-X)}[1]\otimes(f_{2*}j_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})[1])).$$

By [Ka–RLS, 2.9.2], this is (restriction to T of) the asserted middle convolution. QED for 5.3.7

Once we have Proposition 5.3.7, then to prove the irreducibility of $\mathcal{G}$ it suffices to show that $f_{2*}j_{2*}j_1^{\ *}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ is an irreducible middle extension. This is immediate from the Irreducible Induction Criterion 3.3.1, since the singularities of $j_{2*}j_1^{\ *}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ on $C - D_2$ include the $\deg(D_1)$ distinct zeroes of f_1, and the f_2–fibre containing each of these zeroes consists of $\deg(D_2)$ distinct points, precisely one of which, namely that zero, is a singularity of $j_{2*}j_1^{\ *}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$. QED

5.4 Theorems of big monodromy in characteristic not 2

Theorem 5.4.1 Let k be an algebraically closed field of characteristic not 2, C/k a proper, smooth connected curve of genus g. Suppose that $D = \sum a_i P_i$ is an effective divisor of degree $d \geq 2g+1$, with all a_i invertible in k. Let $\mathcal{F}$ be an irreducible middle extension sheaf on C with $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}} := \mathrm{Sing}(\mathcal{F})\cap(C-D)$ nonempty. Suppose that either $\mathcal{F}$ is everywhere tame, or that $\mathcal{F}$ is tame at all points of D and that the characteristic p is either zero or a prime $p \geq \mathrm{rank}(\mathcal{F}) + 2$. Suppose that the following inequalities hold:

if $\mathrm{rank}(\mathcal{F}) = 1$, $\quad 2g-2+d \geq \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, 4\mathrm{rank}(\mathcal{F}))$, if $\mathrm{rank}(\mathcal{F}) \geq 2$, $\quad 2g-2+d \geq \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, 72\mathrm{rank}(\mathcal{F}))$.

Fix a nontrivial character χ whose finite order $n \geq 2$ is invertible in k. Form the lisse sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$$

on the space $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$.

Pick a function f in $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ which also lies in the dense open set U of Theorem 2.2.6 applied with S taken to be $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. Thus f as map from C–D to $\mathbb{A}^1$ is of Lefschetz type, and for each s in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$, the fibre $f^{-1}(s)$ consists of d distinct points, only one of which lies in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. Consider the lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{H}$ on $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$ given by

$$\mathcal{H} := [t \mapsto t-f]^*\mathcal{G},$$

i.e., by

$$t \mapsto H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)}).$$

Its geometric monodromy group G_{geom} is either Sp or SO or O, or G_{geom} contains SL. If $\mathcal{F}$ is orthogonally (respectively symplectically) self–dual, and χ has order 2, then G_{geom} is Sp (respectively SO or O). If χ has order ≥ 3, then G_{geom} contains SL.

proof Let us put $r := \mathrm{rank}(\mathcal{F})$, $m := \#\,\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. We have seen (5.3.5) that $\mathcal{H}$ is the restriction to $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$ of the middle additive convolution of $f_*\mathcal{F}$ and $\mathcal{L}_\chi$.

Let us put

$$\mathcal{F}_1 := f_*\mathcal{F}.$$

We have seen above (in the proof of 5.3.6) that $\mathcal{F}_1$ is an irreducible middle extension on $\mathbb{A}^1$. Notice that $\mathcal{F}_1$ lies in the class $\mathcal{P}_{\mathrm{conv}}$, cf. 4.0.2, because its rank is ≥ 3. [Indeed, its rank is $d\times\mathrm{rank}(\mathcal{F}) \geq d$. If $g > 0$, then the hypothesis that $d \geq 2g+1$ gives $d \geq 3$. If $g = 0$, the hypothesis

$$2g-2+d \geq \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}, 4\mathrm{rank}(\mathcal{F})).$$

gives $d \geq 6$.]

The sheaf $\mathcal{F}_1$ is tame at ∞, because $\mathcal{F}$ is tame at all the poles of f, and the poles of f all have order prime to p. Moreover, the $I(\infty)$–invariants are given by

$$\mathcal{F}_1(\infty)^{I(\infty)} \cong \oplus_{\text{points } P_i \text{ in } D}\, \mathcal{F}(P_i)^{I(P_i)}.$$

Over each critical value α of f, $\mathcal{F}$ is lisse, and $f-\alpha$ has one and only one double zero, so the local monodromy of $\mathcal{F}_1$ at α is quadratic of drop r, with scale the unique character of order 2:

$$\mathcal{F}_1(\alpha)/\mathcal{F}_1(\alpha)^{I(\alpha)} \cong \text{r copies of} \mathcal{L}_{\chi_2(x-\alpha)}.$$

Over the m images $\delta = f(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$, f is finite etale, and β is the unique point of $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ in the fibre, so the local monodromy of $\mathcal{F}_1$ at δ has drop $\leq$ r. More precisely, we have

$$\mathcal{F}_1(\delta)/\mathcal{F}_1(\delta)^{I(\delta)} \cong \mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)},$$

where we use f to identify $I(\delta)$ with $I(\beta)$.

At all other points of $\mathbb{A}^1$, i.e., on $\mathbb{A}^1$ – CritVal(f, $\mathcal{F}$), $\mathcal{F}_1$ is lisse. Moreover, if $\mathcal{F}$ is everywhere tame on C, then $\mathcal{F}_1$ is everywhere tame. Now form $\mathcal{H}$, the middle additive convolution of $\mathcal{F}_1$ with $\mathcal{L}_\chi$. Thus $\mathcal{H}$ is tame at ∞ (by 4.1.10, part 2d)), and it is everywhere tame if $\mathcal{F}$ is everywhere tame (by 4.1.10, parts 1b) and 2d)). By 5.2.1, part 6), we have the inequality

$$\mathrm{rank}(\mathcal{H}) \geq (2g-2+d)r + \#\mathrm{Sing}_{\mathrm{finite}}(\mathcal{F}) > (2g-2+d)r.$$

The local monodromy of $\mathcal{H}$ at each of the m images $\delta = f(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ has drop $\leq$ r, by 4.1.10, part 1c), and is given by

$$\mathcal{H}(\delta)/\mathcal{H}(\delta)^{I(\delta)} \cong \mathrm{MC}_\chi \mathrm{loc}(\delta)(\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)} \text{ as } I(\delta)\text{–rep'n}.$$

The local monodromy of $\mathcal{H}$ at each critical value α of f is quadratic of drop r, with scale the character $\chi\chi_2$:

$$\mathcal{H}(\alpha)/\mathcal{H}(\alpha)^{I(\alpha)} \cong \mathcal{L}_{\chi(x-\alpha)} \otimes (\text{r copies of } \mathcal{L}_{\chi_2(x-\alpha)})$$
$$\cong \text{r copies of } \mathcal{L}_{\chi\chi_2(x-\alpha)}.$$

The key observation here is that $\chi\chi_2$ is **not** of order two, and that f **has** critical points (because their number, the number of zeroes of df, is

$$2g-2+ \Sigma_i(1+a_i) > 2g-2+d > 2\#\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}} > 2 > 0).$$

The conclusion follows from Theorem 1.5.1 (and Theorem 1.7.1, if r=1), applied to the data (r, m, $\mathcal{H}$). QED

Proposition 5.4.2 Hypotheses and notations as in Theorem 5.4.1 above, suppose that χ has order 2, but $\mathcal{F}$ is not self–dual. Then G_{geom} contains SL.

proof If not, then by the paucity of choice, G_{geom} is contained in either Sp or O, and hence $\mathcal{H}$ is self–dual. But $\mathcal{H}$ is the middle convolution of $f_*\mathcal{F}$ and $\mathcal{L}_\chi$. As χ has order 2, we recover $f_*\mathcal{F}$ as the middle convolution of $\mathcal{H}$ and $\mathcal{L}_\chi$. As χ has order 2, $\mathcal{L}_\chi$ is self–dual. As both $\mathcal{H}$ and $\mathcal{L}_\chi$ are self–dual, so is their middle convolution, $f_*\mathcal{F}$. By Proposition 3.4.1, the autoduality of $f_*\mathcal{F}$ implies that of $\mathcal{F}$, contradiction. QED

Proposition 5.4.3 Hypotheses and notations as in Theorem 5.4.1 above, suppose that χ has order 2, and that $\mathcal{F}$ is symplectically self–dual.

1) Suppose there exists a finite singularity β of $\mathcal{F}$, i.e., a point β in $\mathrm{Sing}(\mathcal{F}) \cap (C-D)$, such that the following two conditions hold:

1a) $\mathcal{F}$ is tame at β,

1b) $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)}$ has odd dimension.

Then the group $\mathrm{G}_{\mathrm{geom}}$ for the sheaf $\mathcal{H}$ is the full orthogonal group O.

2) Suppose that $\mathcal{F}$ is everywhere tame. Then $\mathrm{G}_{\mathrm{geom}}$ for $\mathcal{H}$ is the special orthogonal group SO if and only if $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)}$ has even dimension for every finite singularity β of $\mathcal{F}$.

proof In terms of $\mathcal{F}_1 := \mathrm{f}_*\mathcal{F}$, $\mathcal{H}$ on $\mathbb{A}^1 - \mathrm{CritVal}(\mathrm{f}, \mathcal{F})$ is (the restriction from $\mathbb{A}^1$ of) the middle convolution $\mathcal{F}_1 *_{\mathrm{mid}+} \mathcal{L}_\chi$. We already know that $\mathrm{G}_{\mathrm{geom}}$ for $\mathcal{H}$ is either SO or O, so we have only to see whether $\det(\mathcal{H})$ is trivial or not. Since $\det(\mathcal{H})$ is either trivial or of order 2, it is **tame** on $\mathbb{A}^1 - \mathrm{CritVal}(\mathrm{f}, \mathcal{F})$. Hence $\det(\mathcal{H})$ is trivial if and only if it is trivial on every **finite** inertia group $\mathrm{I}(\gamma)$, γ in $\mathrm{CritVal}(\mathrm{f}, \mathcal{F})$.

At γ which is a critical value α of f, we have seen that the local monodromy of $\mathcal{F}_1$ at α is quadratic of drop r := rank($\mathcal{F}$), with scale the unique character of order 2:

$$\mathcal{F}_1(\alpha)/\mathcal{F}_1(\alpha)^{\mathrm{I}(\alpha)} \cong \text{r copies of} \mathcal{L}_{\chi_2(\mathrm{x}-\alpha)}.$$

The local monodromy of $\mathcal{H} = \mathcal{F}_1 *_{\mathrm{mid}+} \mathcal{L}_\chi$ at α is given by

$$\begin{aligned}\mathcal{H}(\alpha)/\mathcal{H}(\alpha)^{\mathrm{I}(\alpha)} &\cong (\mathcal{F}_1(\alpha)/\mathcal{F}_1(\alpha)^{\mathrm{I}(\alpha)})\otimes\mathcal{L}_{\chi(\mathrm{x}-\alpha)} \\ &\cong \text{r copies of } \mathbb{1},\end{aligned}$$

this last equality because χ is the quadratic character χ_2. From this we calculate

$$\det(\mathcal{H}(\alpha)) = \det(\mathcal{H}(\alpha)/\mathcal{H}(\alpha)^{\mathrm{I}(\alpha)}) = \mathbb{1}.$$

Thus the local monodromy of $\det(\mathcal{H})$ is trivial at all the critical values of f.

At γ which is the image $\delta = \mathrm{f}(\beta)$ of a point β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$, we have seen that

$$\mathcal{F}_1(\delta)/\mathcal{F}_1(\delta)^{\mathrm{I}(\delta)} \cong \mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)}$$

where we use f to identify $\mathrm{I}(\delta)$ with $\mathrm{I}(\beta)$. Using this identification, the local monodromy of $\mathcal{H} = \mathcal{F}_1 *_{\mathrm{mid}+} \mathcal{L}_\chi$ at δ is

$$\mathcal{H}(\delta)/\mathcal{H}(\delta)^{\mathrm{I}(\delta)} \cong \mathrm{MC}_\chi\mathrm{loc}(\delta)(\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)}) \text{ as } \mathrm{I}(\delta)\text{-rep'n}.$$

If $\mathcal{F}$ is tame at β, we have

$$\mathcal{H}(\delta)/\mathcal{H}(\delta)^{\mathrm{I}(\delta)} \cong (\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)})\otimes\mathcal{L}_{\chi(\mathrm{x}-\delta)}.$$

We then readily compute

$$\begin{aligned}\det(\mathcal{H}(\delta)) &= \det(\mathcal{H}(\delta)/\mathcal{H}(\delta)^{\mathrm{I}(\delta)}) \\ &= \det((\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)})\otimes\mathcal{L}_{\chi(\mathrm{x}-\delta)}) \\ &= \det((\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)}))\otimes(\mathcal{L}_{\chi(\mathrm{x}-\delta)})^{\dim(\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)})}.\end{aligned}$$

But we have

$$\det((\mathcal{F}(\beta)/\mathcal{F}(\beta)^{\mathrm{I}(\beta)})) = \det(\mathcal{F}(\beta)) = \mathbb{1},$$

this last equality because $\mathcal{F}$ is symplectic, and $\mathrm{Sp} \subset \mathrm{SL}$. Thus we find

$$\det(\mathcal{H}(\delta)) = (\mathcal{L}_{\chi(x-\delta)})^{\dim(\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)})}.$$

Thus $\det(\mathcal{H})$ is nontrivial at the image $\delta = f(\beta)$ of a point β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ at which $\mathcal{F}$ is tame, if and only if $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)}$ has odd dimension. This proves 1).

Suppose now that $\mathcal{F}$ is everywhere tame. We already know that $\det(\mathcal{H})$ is trivial at all the critical values of f, so $\det(\mathcal{H})$ is trivial if and only if it is trivial at every $\delta = f(\beta)$, β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. For $\mathcal{F}$ everywhere tame, this triviality at every $\delta = f(\beta)$ means precisely that $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)}$ has even dimension for all finite singularities β of $\mathcal{F}$. QED

Remark 5.4.4 Here is an example to show that part 2) of the above proposition can fail if we drop the hypothesis that $\mathcal{F}$ be everywhere tame. We fix an even integer $2n \geq 2$, and work over $\overline{\mathbb{F}}_p$ for any prime $p \geq 2n+2$. Fix a nontrivial $\overline{\mathbb{Q}}_\ell$–valued additive character ψ of $\mathbb{F}_p$. Denote by Kl_{2n} the standard Kloosterman sheaf in 2n variables: thus Kl_{2n} is the lisse sheaf of rank 2n on $\mathbb{G}_m/\mathbb{F}_p$ whose trace function at a point α in $E^\times$, E a finite extension E of $\mathbb{F}_p$, is

$$\mathrm{Trace}(\mathrm{Frob}_{\alpha,E} \mid \mathrm{Kl}_{2n}) = -\Sigma_{x_1 x_2 \ldots x_{2n} = \alpha \text{ in } E}\, \psi(\Sigma x_i).$$

One knows that Kl_{2n} is symplectically self–dual.

Take $\mathcal{F}$ the middle extension of the lisse sheaf $[x \mapsto 1/x]^*\mathrm{Kl}_{2n}$ on $\mathbb{G}_m$. One knows that $\mathrm{Kl}_{2n}(\infty)$ is a totally wild irreducible representation of $I(\infty)$, all of whose slopes are 1/2n. Thus $\mathcal{F}$ is totally wild at zero, and hence $\mathcal{F}(0)/\mathcal{F}(0)^{I(0)}$ has even dimension 2n.

We take C to be $\mathbb{P}^1/\overline{\mathbb{F}}_p$, D to be $d\infty$ for a sufficiently large integer d prime to p, χ to be the quadratic character χ_2, and $\mathcal{F}$ as above. Then $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ is $\{0\}$, and, as noted above, $\mathcal{F}(0)/\mathcal{F}(0)^{I(0)}$ has even dimension 2n. Nonetheless, we will see that G_{geom} for $\mathcal{H}$ is the full orthogonal group O. More precisely, with $\delta := f(0)$, we will show that $\det(\mathcal{H})$ is nontrivial at δ. To simplify the notations, let us replace f by $f - \delta$, so that $f(0) = 0$. Then we have

$$\mathcal{H}(0)/\mathcal{H}(0)^{I(0)} \cong \mathrm{MC}_\chi\mathrm{loc}(0)(\mathcal{F}(0)/\mathcal{F}(0)^{I(0)}).$$

We will show that $\det(\mathcal{H}(0))$ is $\mathcal{L}_\chi$. We have

$$\det(\mathcal{H}(0)) = \det(\mathcal{H}(0)/\mathcal{H}(0)^{I(0)}) = \det(\mathrm{MC}_\chi\mathrm{loc}(0)(\mathcal{F}(0)/\mathcal{F}(0)^{I(0)})).$$

We will calculate $\mathrm{MC}_\chi\mathrm{loc}(0)(\mathcal{F}(0)/\mathcal{F}(0)^{I(0)})$ by a global argument. The sheaf $\mathcal{F}$ on $\mathbb{A}^1$ lies in $\mathcal{P}_{\mathrm{conv}}$ of 4.0.2. We define

$$\mathcal{G} := \mathcal{F} *_{\mathrm{mid}+} \mathcal{L}_\chi \text{ in } \mathcal{P}_{\mathrm{conv}}.$$

Then by Theorem 4.1.10, part 1) we have

$$\mathcal{G}(0)/\mathcal{G}(0)^{I(0)} \cong \mathrm{MC}_\chi\mathrm{loc}(0)(\mathcal{F}(0)/\mathcal{F}(0)^{I(0)}).$$

Thus

$$\det(\mathcal{H}(0)) = \det(\mathrm{MC}_\chi\mathrm{loc}(0)(\mathcal{F}(0)/\mathcal{F}(0)^{\mathrm{I}(0)}))$$
$$= \det(\mathcal{G}(0)/\mathcal{G}(0)^{\mathrm{I}(0)})$$
$$= \det(\mathcal{G}(0)).$$

Hence we are reduced to showing that $\det(\mathcal{G}(0))$ is $\mathcal{L}_\chi$.

Applying Fourier transform FT ($:= \mathrm{FT}_\psi$) to the defining equation
$$\mathcal{G} := \mathcal{F}*_{\mathrm{mid}+}\mathcal{L}_\chi,$$
we obtain
$$\mathrm{FT}(\mathcal{G}) = j_*(\mathrm{FT}(\mathcal{F})\otimes\mathcal{L}_\chi \mid \mathbb{G}_m).$$

The key observation is that, because $\mathcal{F}$ is $[x \mapsto 1/x]^*\mathrm{Kl}_{2n}$, we have
$$\mathrm{FT}(\mathcal{F}) \cong \mathrm{Kl}_{2n+1},$$
a remark due to Deligne [De–AFT, 7.1.4] and developed in [Ka–ESDE, 8.1.12 and 8.4.3]. Thus we find
$$\mathrm{FT}(\mathcal{G}) = j_*(\mathrm{FT}(\mathcal{F})\otimes\mathcal{L}_\chi \mid \mathbb{G}_m) = j_!\mathrm{Kl}_{2n+1}(\chi, \chi, ..., \chi).$$

We can calculate $\mathrm{FT}(j_!\mathrm{Kl}_{2n+1}(\chi, \chi, ..., \chi))$ as a hypergeometric sheaf of type (1, 2n+1), cf. [Ka–ESDE, 9.3.2 with d=1]. The result is
$$\mathrm{FT}(j_!\mathrm{Kl}_{2n+1}(\chi, \chi, ..., \chi)) \cong j_*\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi).$$

Since FT is involutive, we find a geometric isomorphism
$$[x \mapsto -x]^*\mathcal{G} \cong j_*\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi).$$

So to show that $\det(\mathcal{G}(0))$ is $\mathcal{L}_\chi$, it is equivalent to show that $\det(\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi))(0)$ is $\mathcal{L}_\chi$.

The sheaf $\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi)$ is lisse on $\mathbb{G}_m$. Its local monodromy at ∞ is $\mathcal{L}_\chi\otimes\mathrm{Unip}(2n+1)$, whose determinant is $\mathcal{L}_\chi$ (remember χ is χ_2). Its local monodromy at 0 is $\mathbb{1}\oplus W$, where W has rank 2n and all slopes 1/2n. Since all slopes at 0 are < 1, $\det(\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi))$ is tame at 0. Thus $\det(\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi))$ is lisse on $\mathbb{G}_m$, tame at both 0 and ∞, and agrees with $\mathcal{L}_\chi$ at ∞. Therefore we have a global isomorphism
$$\det(\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi)) \cong \mathcal{L}_\chi \text{ on } \mathbb{G}_m/\overline{\mathbb{F}}_p.$$

In particular, $\det(\mathcal{H}yp(\mathbb{1}; \chi, ..., \chi))(0)$ is $\mathcal{L}_\chi$.

Here is a further elaboration on this sort of counterexample. With 2n, p, and d fixed as above, choose further an **odd** integer $k \geq 1$ which is prime to p. Now define $\mathcal{F}$ to be the middle extension of the lisse sheaf $[x \mapsto 1/x^k]^*\mathrm{Kl}_{2n}$ on $\mathbb{G}_m$. Then $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ is $\{0\}$, $\mathcal{F}$ is totally wild at 0, and $\mathcal{F}(0)/\mathcal{F}(0)^{\mathrm{I}(0)}$ has even dimension 2n. Using [Ka–ESDE, 9.3.2 with d=k], a similar argument now shows that $\mathcal{H}$ has G_{geom} the full orthogonal group, and that $\det(\mathcal{H})$ is nontrivial at 0.

(5.4.5) We will now give another one–parameter family of twists with big monodromy. Before stating the result, we need an elementary lemma.

Lemma 5.4.6 Let k be an algebraically closed field of any characteristic, C/k a proper, smooth connected curve of genus g. Suppose that $D = \sum a_i P_i$ is an effective divisor of degree d. Suppose d_1 and d_2 are positive integers with $d_1 + d_2 = d$. If k has characteristic $p > 0$, suppose further that $d_2/d \le (p-1)/p$. Then we can write D as a sum of effective divisors $D_1 + D_2$ with D_2 of degree either d_2 or $d_2 + 1$, such that $D_2 = \sum c_i P_i$, has all its nonzero c_i invertible in k.

proof If k has characteristic zero, any writing of D as a sum of effective divisors $D_1 + D_2$ with D_2 of degree d_2 does the job.

If k has characteristic $p > 0$, put $\lambda := d_2/d$. For real $x \ge 0$, we denote its "floor" and "ceiling"

$$[x]_{fl} := \text{the greatest integer} \le x,$$
$$[x]_{ce} := \text{the least integer} \ge x.$$

Since $\lambda \le 1$, we have, for each i,

$$a_i \ge [\lambda a_i]_{ce} \ge \lambda a_i \ge [\lambda a_i]_{fl}.$$

We define effective divisors D_{fl} and D_{ce} by

$$D_{fl} := \sum_i [\lambda a_i]_{fl} P_i,\ D_{ce} := \sum_i [\lambda a_i]_{ce} P_i.$$

Thus $D \ge D_{ce} \ge D_{fl}$, and $\deg(D_{ce}) \ge d_2 \ge \deg(D_{fl})$. For each i, the coefficients $[\lambda a_i]_{ce}$ and $[\lambda a_i]_{fl}$ are either equal or differ by 1. So we can choose, for each i, either $[\lambda a_i]_{ce}$ and $[\lambda a_i]_{fl}$, call it b_i, so that the "intermediate" divisor $D_{int} := \sum_i b_i P_i$ has degree d_2. Clearly

$$D_{ce} \ge D_{int} \ge D_{fl}.$$

If D_{int} has all its nonzero b_i invertible in k, we take D_2 to be D_{int}. Then D_2 will have degree d_2.

If some of the nonzero b_i are divisible by p, we modify D_{int} as follows. First of all, if p divides a nonzero b_i, then $b_i \ge p$, so $b_i - 1$ is positive and prime to p. What about $b_i + 1$? It is prime to p, but is $b_i + 1 \le a_i$? In other words, is $b_i < a_i$? The answer is yes, because if not, then $b_i = a_i$. But $a_i \ge [\lambda a_i]_{ce} \ge b_i$, so we would have $a_i = [\lambda a_i]_{ce}$. This means in turn that $\lambda a_i > a_i - 1$, i.e., $1 > a_i(1-\lambda)$. But p divides b_i, so $a_i \ge p$, and so $1 > p(1-\lambda)$, which contradicts the hypothesis $\lambda \le (p-1)/p$.

So each nonzero b_i that is divisible by p can be either increased by 1 or decreased by 1 and continue to lie in the range $[0, a_i]$. If there are evenly many indices i whose b_i is divisible by p, increase half of them by 1 and decrease the other half by 1, to get the desired D_2: it has degree d_2. If there are oddly many b_i divisible by p, group all but one in pairs, and in each pair increase one member by 1 and decrease the other by 1. Increase the leftover by 1. This gives a D_2 of degree $1+d_2$. QED

Remark 5.4.7 The example of a divisor D of the form $\sum_i pP_i$, which has all its $a_i = p$, shows that the hypothesis $d_2/d \le (p-1)/p$ cannot be relaxed. The example of a divisor D of the form dP, and the choice $d_2 = p$, shows that we cannot insist that D_2 have degree d_2.

Corollary 5.4.8 Let k be an algebraically closed field, C/k a proper, smooth connected curve of genus g.
1) Suppose that $D = \sum a_i P_i$ is an effective divisor of degree $d \ge 4g+5$. Then we can write D as a sum of effective divisors $D_1 + D_2$ with degrees $d_1 \ge 2g+2$ and $d_2 \ge 2g+2$, such that $D_2 = \sum c_i P_i$ has all its nonzero c_i invertible in k.
2) Suppose that $D = \sum a_i P_i$ is an effective divisor of degree $d \ge 4g+4$. Then we can write D as a sum of effective divisors $D_1 + D_2$ with degrees $d_1 \ge 2g+2$ and $d_2 \ge 2g+1$, such that $D_2 = \sum c_i P_i$ has all its nonzero c_i invertible in k.
3) Fix an integer $A \ge 0$. Suppose that $D = \sum a_i P_i$ is an effective divisor of degree
$$d \ge \mathrm{Max}(6g+9, 6A + 11),$$
and that the characteristic is not two. Then we can write D as a sum of effective divisors $D_1 + D_2$ both of whose degrees d_1 and d_2 are at least 2g+2, such that $D_2 = \sum c_i P_i$ has all its nonzero c_i invertible in k, and such that $2g - 2 + d > 2(A+d_1)$.

proof Assertion 1) is immediate from the lemma, with initial choice $d_2 = 2g+2$.

For 2), write D as the sum of effective divisors E + F with E of degree $e = 4g+2$, and F of degree $f \ge 2$. Apply the lemma to E and the initial choice $d_2 := [e/2]$. Then we end up with E_2 of degree either $[e/2]$ or $[e/2]+1$ (both of which are $\ge [e/2] = 2g+1$), and E_1 of degree either $e - [e/2]$ or $e -1 - [e/2]$ (both of which are $\ge [e/2] - 1 = 2g$). Then $D_1 := E_1 + F$, $D_2 := E_2$, is the desired decomposition.

For 3), we apply the lemma with the initial choice $d_2 := [2d/3]$, allowed because the characteristic is not two. We end up with D_2 of degree d_2 either $[2d/3]$ or $[2d/3]+1$, both of which are $\ge (2d-2)/3$ and both of which are $\le (2d+3)/3$. Then D_1 has degree d_1 either $d - [2d/3]$ or $d - 1 - [2d/3]$, both of which are $\ge (d-3)/3$, and both of which are $\le (d+2)/3$. So both D_1 and D_2 have degree at least $(d-3)/3 \ge 2g+2$. We also have
$$\begin{aligned} 2g - 2 + d - 2(A+d_1) &= 2g - 2 + d_1 + d_2 - 2(A+d_1) \\ &= d_2 - d_1 + 2g - 2 - 2A \\ &\ge d_2 - d_1 - 2A - 2 \end{aligned}$$

$$\geq (2d-2)/3 - (d+2)/3 - 2A - 2$$
$$= (d-4)/3 - 2A - 2$$
$$\geq (6A+7)/3 - 2A - 2 > 0,$$

as required. QED

Theorem 5.4.9 Let k be an algebraically closed field of characteristic not 2, C/k a proper, smooth connected curve of genus g. Suppose that $D = \sum a_i P_i$ is an effective divisor of degree $d \geq 4g+4$. Write D as a sum of effective divisors $D_1 + D_2$ of degrees $d_1 \geq 2g+2$ and $d_2 \geq 2g+1$, such that $D_2 = \sum c_i P_i$ has all its nonzero c_i invertible in k.

Let $\mathcal{F}$ be an irreducible middle extension sheaf on C. Suppose that either $\mathcal{F}$ is everywhere tame, or that $\mathcal{F}$ is tame at all points of D and that the characteristic p is either zero or a prime $p \geq \mathrm{rank}(\mathcal{F}) + 2$. Suppose that the following inequalities hold:

if $\mathrm{rank}(\mathcal{F}) = 1$, $2g - 2 + d > \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F}) \cap (C-D_2)), 4\mathrm{rank}(\mathcal{F}))$,
if $\mathrm{rank}(\mathcal{F}) \geq 2$, $2g - 2 + d > \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F}) \cap (C-D_2)), 72\mathrm{rank}(\mathcal{F}))$.

Fix a nontrivial character χ of finite order $n \geq 2$. If n=4, suppose also that
$$2g - 2 + d > 2(\#(\mathrm{Sing}(\mathcal{F}) \cap (C-D_2)) + d_1).$$
If n=4 and the curve C has genus g=0, suppose in addition that D_1 and D_2 are chosen so that $d_2 \geq 2$. (Such a choice is always possible if g=0 by Corollary 5.8.4, part 1), because $d-2 = 2g-2+d > 72\mathrm{rank}(\mathcal{F}) \geq 72$, hence $d \geq 75 > 4g+5$.)

Fix a function
$$f_1 \text{ in } \mathrm{Fct}(C, \deg(D_1), D_1, \mathrm{Sing}(\mathcal{F}) \cup D^{\mathrm{red}}).$$
Fix a function f_2 in $\mathrm{Fct}(C, \deg(D_2), D_2, \mathrm{Sing}(\mathcal{F}) \cup D^{\mathrm{red}} \cup f_1^{-1}(0))$ which also lies in the open set U of Theorem 2.2.6 with respect to the set $S := f_1^{-1}(0) \cup (\mathrm{Sing}(\mathcal{F}) \cap (C-D_2))$. Consider the lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{H}$ on $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F} \otimes \mathcal{L}_{\chi(f_1)})$ given by $[t \mapsto f_1(t-f_2)]^*\mathcal{G}$, i.e., by
$$t \mapsto H^1(C, j_*(\mathcal{F} \otimes \mathcal{L}_{\chi(f_1(t-f_2))}).$$
Its geometric monodromy group G_{geom} is either Sp or SO or O or a group between SL and GL. If $\mathcal{F}$ is orthogonally (respectively symplectically) self–dual, and χ has order 2, then G_{geom} is Sp (respectively SO or O). If χ has order ≥ 3, then G_{geom} contains SL.

proof Suppose first $n \neq 4$. Put $r := \mathrm{rank}(\mathcal{F})$, $m := \#(\mathrm{Sing}(\mathcal{F}) \cap (C-D_2))$. We have seen in Proposition 5.3.7 that $\mathcal{H}$ is the restriction to $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F} \otimes \mathcal{L}_{\chi(f_1)})$ of the middle additive convolution of $f_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ and $\mathcal{L}_\chi$.

Let us put

$$\mathcal{F}_1 := f_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}).$$

As already noted at the end of the proof of 5.3.6, the Irreducible Induction Criterion 3.3.1 shows that $\mathcal{F}_1$ is an irreducible middle extension sheaf. The sheaf $\mathcal{F}_1$ lies in the class $\mathcal{P}_{\mathrm{conv}}$, because it has at least $d_1 \geq 2g+2 \geq 2$ finite singularities, namely the d_1 distinct images by f_2 of the d_1 distinct zeroes of f_1. It is tame at ∞, because $\mathcal{F}$ is tame at all the poles of f_2, and the poles of f_2 all have order prime to p.

Over each critical value α of f_2, $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$ is lisse, and $f_2-\alpha$ has one and only one double zero, so the local monodromy of $\mathcal{F}_1$ at α is quadratic of drop r, with scale the unique character of order 2:

$$\mathcal{F}_1(\alpha)/\mathcal{F}_1(\alpha)^{I(\alpha)} \cong \text{r copies of} \mathcal{L}_{\chi_2(x-\alpha)}.$$

Over the m images $\delta = f_2(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})\cap(C-D_2)$, f_2 is finite etale, and β is the unique point of $\mathrm{Sing}(\mathcal{F})\cap(C-D_2)$ in the fibre, so the local monodromy of $\mathcal{F}_1$ at δ has drop $\leq$ r. More precisely, we have

$$\mathcal{F}_1(\delta)/\mathcal{F}_1(\delta)^{I(\delta)} \cong \mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)},$$

where we use f_2 to identify $I(\delta)$ with $I(\beta)$.

Over each of the d_1 images $\gamma = f_2(\zeta)$ of the zeroes of f_1, f_2 is finite etale, ζ is the only zero of f_1 in its f_2–fibre, and $\mathcal{F}$ is lisse. Thus $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$ is lisse at all but the point ζ in the fibre $f_2^{-1}(\gamma)$. At ζ the local monodomy of $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$ is quadratic of drop r, with scale the character $\mathcal{L}_{\chi(\text{uniformizer at } \zeta)}$ of $I(\zeta)$. Thus the local monodromy of $\mathcal{F}_1$ at γ is quadratic of drop r, with scale the character $\mathcal{L}_{\chi(x-\gamma)}$ of $I(\gamma)$.

At all other points of $\mathbb{A}^1$, i.e., on $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$, $\mathcal{F}_1$ is lisse. Moreover, if $\mathcal{F}$ is everywhere tame on C, then $\mathcal{F}_1$ is everywhere tame. Now form $\mathcal{H}$, the middle additive convolution of $\mathcal{F}_1$ with $\mathcal{L}_\chi$:

$$\mathcal{H} := \mathcal{F}_1 *_{\mathrm{mid}+} \mathcal{L}_\chi.$$

Thus (by 4.1.10, 2d) and 1b)) $\mathcal{H}$ is tame at ∞, and it is everywhere tame if $\mathcal{F}$ is everywhere tame. Its rank is given by (5.2.1, part 5))

$$\begin{aligned}\mathrm{rank}(\mathcal{H}) = {} & (2g-2 + d)r \\ & + \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}) \\ & + \textstyle\sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{drop}_s(\mathcal{F}) \\ & + \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{drop}_{P_i}(\mathcal{F}(P_i)\otimes(\mathcal{L}_\chi a_i)(\infty, P_i)),\end{aligned}$$

where we have written $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ for $\mathrm{Sing}(\mathcal{F})\cap(\mathrm{C-D})$.
In particular, we have the inequality (5.2.1, part 6))

$$\mathrm{rank}(\mathcal{H}) \geq (2g-2+d)r.$$

The local monodromy of $\mathcal{H}$ at the m images $\delta = f_2(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})\cap(\mathrm{C-D}_2)$ has drop $\leq$ r, by (4.1.10, part 1c), applied to $\mathcal{F}_1$).

The local monodromy of $\mathcal{H}$ at each critical value α of f_2 is quadratic of drop r, with scale the character $\chi\chi_2$:

$$\mathcal{H}(\alpha)/\mathcal{H}(\alpha)^{\mathrm{I}(\alpha)} \cong \mathcal{L}_{\chi(x-\alpha)}\otimes(\text{r copies of } \mathcal{L}_{\chi_2(x-\alpha)})$$
$$\cong \text{r copies of } \mathcal{L}_{\chi\chi_2(x-\alpha)}.$$

Over each of the d_1 images $\gamma = f_2(\zeta)$ of the zeroes of f_1, the local monodromy of $\mathcal{H}$ at γ is quadratic of drop r, with scale the character $\mathcal{L}_{\chi^2(x-\gamma)}$ of $\mathrm{I}(\gamma)$.

With the exception of at most m points of $\mathbb{A}^1$, namely the images by f_2 of points in $\mathrm{Sing}(\mathcal{F})\cap(\mathrm{C-D}_2)$, the local monodromy of $\mathcal{H}$ is quadratic of drop r, with scale a character not of order 2. Indeed, at the critical values of f_2, $\chi\chi_2$ is not of order 2 (χ being nontrivial), and at the d_1 images of the zeroes of f_1, χ^2 is not of order 2 (because the order n of χ is assumed to be not 4).

The conclusion now follows from Theorem 1.5.1 (and Theorem 1.7.1, if r=1), applied to the data (r, m, $\mathcal{H}$).

Suppose now that n is 4. Our $\mathcal{F}_1$ is still perverse irreducible, and in the class $\mathcal{P}_{\mathrm{conv}}$. The difficulty with the case n=4 is this: at the d_1 images $\gamma = f_2(\zeta)$ of the zeroes of f_1, the local monodromy of $\mathcal{H}$ at γ is quadratic of drop r, with scale the character $\mathcal{L}_{\chi^2(x-\gamma)}$ of $\mathrm{I}(\gamma)$. But for χ of order 4, χ^2 is the quadratic character, and so these d_1 points will be part of the excluded "at all but at most m points" in hypothesis 4) of Theorem 1.5.1. To overcome this difficulty, we assume that

$$2g - 2 + d > 2(\#(\mathrm{Sing}(\mathcal{F})\cap(\mathrm{C-D}_2)) + d_1).$$

We put $r := \mathrm{rank}(\mathcal{F})$, $m := \#(\mathrm{Sing}(\mathcal{F})\cap(\mathrm{C-D}_2)) + d_1$. We have noted above that

$$\mathrm{rank}(\mathcal{H}) \geq (2g-2+d)r,$$

so we have

$$\mathrm{rank}(\mathcal{H}) > \mathrm{Max}(2mr, 72r^2).$$

With the exception of at most m points of $\mathbb{A}^1$, namely the images by f_2 of points in $\mathrm{Sing}(\mathcal{F})\cap(\mathrm{C-D}_2)$ and the d_1 images by f_2 of the zeroes of f_1, the local monodromy of $\mathcal{H}$ is quadratic of drop r, with scale a character not of order 2 (in fact, of order 4). The key point is that the remaining finite singularities of $\mathcal{H}$ are at the critical values of f_2, where the local monodromy is quadratic of drop r, with scale $\chi\chi_2$, which has order 4. [The number of critical values is

$$2g-2 + \textstyle\sum_i (1+c_i).$$

This number is strictly positive unless g=0 and $d_2 = 1$. This exceptional case (g=0, d_2=1) is not allowed if n is 4.]

The result now follows from Theorem 1.5.1 (and Theorem 1.7.1, if r=1), applied to the data (r, m, $\mathcal{H}$). QED

Exactly as in Proposition 5.4.2 above, we have

Proposition 5.4.10 Hypotheses and notations as in Theorem 5.4.9 above, suppose that χ has order 2, but $\mathcal{F}$ is not self–dual. Then G_{geom} contains SL.

proof If not, then exactly as in the proof of Proposition 5.4.2, we infer that $f_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ is self–dual, and then that $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$, and hence $\mathcal{F}$, are self–dual. QED

Proposition 5.4.11 Hypotheses and notations as in Theorem 5.4.9 above, suppose that χ has order 2, and that $\mathcal{F}$ is symplectically self–dual.

1) Suppose there exists a D_2–finite singularity β of $\mathcal{F}$, i.e., a point β in Sing($\mathcal{F}$)$\cap$(C–D_2), such that the following two conditions hold:

1a) $\mathcal{F}$ is tame at β,

1b) $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)}$ has odd dimension.

Then the group G_{geom} for the sheaf $\mathcal{H}$ is the full orthogonal group O.

2) Suppose that $\mathcal{F}$ is everywhere tame. Then G_{geom} for $\mathcal{H}$ is the special orthogonal group SO if and only if $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)}$ has even dimension for every D_2–finite singularity β of $\mathcal{F}$.

proof This is proven by essentially recopying the proof of 5.4.3, applied to the sheaf $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$ and the function f_2 (remember that f_1 is chosen to be invertible at β, so $\mathcal{L}_{\chi(f_1)}$ is lisse at β). QED

5.5 Theorems of big monodromy for $\mathcal{G}$:= Twist$_{\chi,C,D}(\mathcal{F})$ on *Fct*(C, d, D, Sing($\mathcal{F}$)$_{finite}$) in characteristic not 2

Theorem 5.5.1 Let k be an algebraically closed field in which 2 is invertible. Fix a prime number ℓ which is invertible in k. Fix a character χ of finite order $n \geq 2$ of the tame fundamental group of $\mathbb{G}_m$/k. Let C/k be a proper smooth connected curve of genus g. Fix an irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on C. Let D = $\sum a_i P_i$ be an effective divisor of degree d on C. Suppose that either

1a) $d \geq 2g+1$, all a_i are invertible in k, Sing($\mathcal{F}$)$\cap$(C–D) is nonempty, and the following inequalities hold:

if $\mathrm{rank}(\mathcal{F}) = 1$, $2g - 2 + d \geq \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D)), 4\mathrm{rank}(\mathcal{F}))$,
if $\mathrm{rank}(\mathcal{F}) \geq 2$, $2g - 2 + d \geq \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D)), 72\mathrm{rank}(\mathcal{F}))$,

or

1b) $d \geq 4g+4$, the following inequalities hold:
if $\mathrm{rank}(\mathcal{F}) = 1$, $2g - 2 + d > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}), 4\mathrm{rank}(\mathcal{F}))$,
if $\mathrm{rank}(\mathcal{F}) \geq 2$, $2g - 2 + d > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}), 72\mathrm{rank}(\mathcal{F}))$,
and, if $n=4$,
$d \geq \mathrm{Max}(6g+9, 6\#\mathrm{Sing}(\mathcal{F}) + 11)$.

Suppose further that
2) either $\mathcal{F}$ is everywhere tame, or $\mathcal{F}$ is tame at all points of D and the characteristic p is either zero or $p \geq \mathrm{rank}(\mathcal{F}) + 2$.
Then the lisse sheaf $\mathcal{G}$ on $\mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ given by

$$f \mapsto H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$$

has G_{geom} given as follows:
a) If $\mathcal{F}$ is orthogonally self–dual, and χ has order 2, then G_{geom} is Sp.
b) If $\mathcal{F}$ is symplectically self–dual, and χ has order 2, then G_{geom} is either SO or O.
c) If either $\mathcal{F}$ is not self–dual or if χ has order > 2, then G_{geom} contains SL.

proof If χ has order two and $\mathcal{F}$ is orthogonally (respectively symplectically) self–dual, then $\mathcal{G}$ is symplectically (resp. orthogonally) self–dual, and we have a priori inclusions

$$G_{\mathrm{geom}} \subset \mathrm{Sp} \text{ (resp. } G_{\mathrm{geom}} \subset \mathrm{O}).$$

In general, we have an a priori inclusion

$$G_{\mathrm{geom}} \subset \mathrm{GL}.$$

Given a smooth connected curve U/k and a map

$$\pi : U \to \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}),$$

we have an a priori inclusion

$$G_{\mathrm{geom}}(\pi^*\mathcal{G} \text{ on } U) \subset G_{\mathrm{geom}}(\mathcal{G} \text{ on } \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})).$$

So it suffices to produce a π such that $G_{\mathrm{geom}}(\pi^*\mathcal{G}$ on U) contains, in the three cases, the groups Sp, SO, and SL respectively. This is precisely what we have done in Theorem 5.4.1 (under hypotheses 1a) and 2)) and in Theorem 5.4.9 (under hypotheses 1b) and 2)). QED

Proposition 5.5.2 Hypotheses and notations as in Theorem 5.5.1 above, suppose that χ has order 2, and that $\mathcal{F}$ is symplectically self–dual.
1) Suppose that there exists a finite singularity β of $\mathcal{F}$, i.e., a point β in $\mathrm{Sing}(\mathcal{F})\cap(C-D)$, such that the following two conditions hold:
1a) $\mathcal{F}$ is tame at β,
1b) $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)}$ has odd dimension.

Then the group G_{geom} for the sheaf $\mathcal{G}$ is the full orthogonal group O.

2) Suppose we are in case 1b) of Theorem 5.5.1, and that there exists a singularity β of $\mathcal{F}$ (but here we do **not** assume that β lies in C−D) such that the following two conditions hold:

2a) $\mathcal{F}$ is tame at β,

2b) $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)}$ has odd dimension.

Suppose further that we can write D as the sum of two effective divisors $D_1 + D_2$ of degrees $d_1 \geq 2g+2$ and $d_2 \geq 2g+1$, such that $D_2 = \sum c_i P_i$ has all its nonzero c_i invertible in k and such that $\beta \in C - D_2$. Then the group G_{geom} for the sheaf $\mathcal{G}$ is the full orthogonal group O.

3) Suppose that the sheaf $\mathcal{G}$ has odd rank. Then the group G_{geom} for the sheaf $\mathcal{G}$ is the full orthogonal group O.

proof If we are in case 1a) of Theorem 5.5.1, then Assertion 1) results from Propostion 5.4.3. If we are in case 1b) of Theorem 5.5.1, then Assertion 1) is a special case of Assertion 2), thanks to Corollary 5.4.8, part 2). Assertion 2) results from Proposition 5.4.11. For assertion 3), we argue as follows. We know that G_{geom} for $\mathcal{G}$ contains SO and is contained in O. To show that G_{geom} is O, it suffices to find a one−parameter family

$$\pi : \mathbb{G}_m \to \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{finite})$$

such that $\det(\pi^*\mathcal{G})$ is nontrivial on $\mathbb{G}_m$.

Fix **any** f in Fct(C, d, D, $\mathrm{Sing}(\mathcal{F})_{finite}$), and consider the map

$$\pi : \mathbb{G}_m \to \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{finite})$$

defined by

$$t \mapsto tf.$$

Thus $\pi^*\mathcal{G}$ is the lisse sheaf on $\mathbb{G}_m$ given by

$$t \mapsto H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(tf)}) = \mathcal{L}_{\chi(t)}\otimes H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})).$$

If $\mathcal{G}$ has odd rank, then $\pi^*\mathcal{G}$ is the direct sum of an odd number of copies of $\mathcal{L}_{\chi(t)}$, and hence, χ being χ_2, $\det(\pi^*\mathcal{G}) \cong \mathcal{L}_{\chi(t)}$. QED

Question 5.5.3 Outside the cases covered by Proposition 5.5.2, we do not know a general, a priori way to distinguish the SO and O cases. The sheaf $\det(\mathcal{G})$ on $\mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{finite})$ is a character of order dividing 2 of $\pi_1(\mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{finite}))$, or, if we like, an element in

$$H^1(\mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{finite}), \boldsymbol{\mu}_2).$$

What is it?

5.6 Theorems of big monodromy in characteristic 2

Theorem 5.6.1 Let k be an algebraically closed field of characteristic 2, C/k a proper, smooth connected curve of genus g. Suppose that $D = \sum a_i P_i$ is an effective divisor of degree $d \geq 6g+3$, with all a_i odd. Let $\mathcal{F}$ be an irreducible middle extension sheaf on C with $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}} := \mathrm{Sing}(\mathcal{F}) \cap (C-D)$ nonempty. Suppose that $\mathcal{F}$ is everywhere tame. Suppose that the degree d is so large that the following inequalities hold:

$$\text{if } \mathrm{rank}(\mathcal{F}) = 1,\ 2g - 2 + d \geq \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D)),\ 4\mathrm{rank}(\mathcal{F})),$$
$$\text{if } \mathrm{rank}(\mathcal{F}) \geq 2,\ 2g - 2 + d \geq \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D)),\ 72\mathrm{rank}(\mathcal{F})),$$

Fix a nontrivial character χ of odd finite order $n \geq 3$. Pick an f in $\mathrm{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ which also lies in the dense open set U of Theorem 2.4.4 applied with S taken to be $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. Thus f as map from C−D to $\mathbb{A}^1$ is of Lefschetz type, each finite monodromy of $f_*\overline{\mathbb{Q}}_\ell$ is a reflection of Swan conductor 1 (by 2.7.1), and for each s in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$, the fibre $f^{-1}(s)$ consists of d distinct points, only one of which lies in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. Consider the lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{H}$ on $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$ given by

$$\mathcal{H} := [t \mapsto t-f]^*\mathcal{G},$$

i.e., by

$$t \mapsto H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(t-f)})).$$

Its geometric monodromy group G_{geom} contains SL.

proof The argument is quite similar to the one given for Theorem 5.4.1. Thus $r := \mathrm{rank}(\mathcal{F})$, $m := \#\,\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$, $\mathcal{F}_1 := f_*\mathcal{F}$, and $\mathcal{H}$ is the restriction to $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$ of the middle additive convolution of $\mathcal{F}_1$ and $\mathcal{L}_\chi$. We know that the function f has

$$g-1 + \sum(1+a_i)/2 \geq (d+1)/2 - 1 \geq (6g+4)/2 - 1 \geq 1$$

critical points, and as many critical values. Over each critical value α of f, $\mathcal{F}$ is lisse, so the local monodromy of $\mathcal{F}_1$ at α is quadratic of drop r, with scale a character ρ_α of $I(\alpha)$ of order 2 and Swan conductor 1:

$$\mathcal{F}_1(\alpha)/\mathcal{F}_1(\alpha)^{I(\alpha)} \cong \text{r copies of } \rho_\alpha.$$

Over the m images $\delta = f(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$, f is finite etale, and β is the unique point of $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ in the fibre, so the local monodromy of $\mathcal{F}_1$ at δ has drop $\leq$ r. More precisely, we have

$$\mathcal{F}_1(\delta)/\mathcal{F}_1(\delta)^{I(\delta)} \cong \mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)},$$

where we use f to identify $I(\delta)$ with $I(\beta)$.

At all other points of $\mathbb{A}^1$, i.e., on $\mathbb{A}^1 - \mathrm{CritVal}(f, \mathcal{F})$, $\mathcal{F}_1$ is lisse. As $\mathcal{F}$ is everywhere tame on C, $\mathcal{F}_1$ is tame except at the critical values of $\mathcal{F}$. Now form $\mathcal{H}$, the middle additive convolution of $\mathcal{F}_1$ with $\mathcal{L}_\chi$. Thus by 4.1.10, 2d), 1b), and 1c), $\mathcal{H}$ is tame at ∞, it is tame outside the critical values of f, and it is lisse outside ∞, the critical values of f, and the m images $\delta = f(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$. Its rank is given by (5.2.1 part 5))

$$\begin{aligned}\mathrm{rank}(\mathcal{H}) = (2g-2+d)r \\ &+ \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{Swan}_s(\mathcal{F}) \\ &+ \textstyle\sum_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}} \mathrm{drop}_s(\mathcal{F}) \\ &+ \textstyle\sum_{P_i \text{ in } D^{\mathrm{red}}} \mathrm{drop}_{P_i}(\mathcal{F}(P_i)\otimes(\mathcal{L}_{\chi}a_i)(\infty, P_i)).\end{aligned}$$

In particular, we have the inequality (5.2.1, part 6))

$$\mathrm{rank}(\mathcal{H}) \geq (2g-2+d)r + \#\mathrm{Sing}_{\mathrm{finite}}(\mathcal{F}) > (2g-2+d)r.$$

The local monodromy of $\mathcal{H}$ at the m images $\delta = f(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ is tame and has drop $\leq$ r, by 4.1.10, part 1c). It is given by

$$\mathcal{H}(\delta)/\mathcal{H}(\delta)^{I(\delta)} \cong \mathrm{MC}_\chi \mathrm{loc}(\delta)(\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)} \text{ as } I(\delta)\text{-rep'n}).$$

The local monodromy of $\mathcal{H}$ at each critical value α of f is quadratic of drop r, with scale a character $\mathrm{MC}_\chi \mathrm{loc}(\alpha)(\rho_\alpha)$ whose order, twice the order of χ by 4.2.2, is ≥ 6. Thus

$$\mathcal{H}(\alpha)/\mathcal{H}(\alpha)^{I(\alpha)} \cong \text{r copies of a character of order} \geq 6.$$

The conclusion follows from Theorem 1.5.1 (and Theorem 1.7.1 if r=1), applied to (r, m, $\mathcal{H}$), with $S - S_0$ the critical values of f, and S_0 the m images $\delta = f(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$.QED

Theorem 5.6.2 Let k be an algebraically closed field of characteristic 2, C/k a proper, smooth connected curve of genus g. Suppose that $D = \sum a_i P_i$ is an effective divisor of degree $d \geq 12g+7$. Write D as a sum of effective divisors $D_1 + D_2$ both of whose degrees d_1 and d_2 are at least 6g+3, such that $D_2 = \sum c_i P_i$ has all its nonzero c_i odd. Let $\mathcal{F}$ be an irreducible middle extension sheaf on C. Suppose that $\mathcal{F}$ is everywhere tame. Suppose that the following inequalities hold:

if $\mathrm{rank}(\mathcal{F}) = 1$, $2g-2+d \geq \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D_2)), 4\mathrm{rank}(\mathcal{F}))$,

if $\mathrm{rank}(\mathcal{F}) \geq 2$, $2g-2+d > \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D_2)), 72\mathrm{rank}(\mathcal{F}))$.

Fix a nontrivial character χ of odd finite order $n \geq 3$.

Fix a function

$$f_1 \text{ in } \mathrm{Fct}(C, \deg(D_1), D_1, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}}).$$

Fix a function f_2 in $\mathrm{Fct}(C, \deg(D_2), D_2, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}}\cup f_1^{-1}(0))$ which also lies in the open set U of Theorem 2.4.4 with respect to the set $S := f_1^{-1}(0)\cup(\mathrm{Sing}(\mathcal{F})\cap(C-D_2))$. Consider the lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{H}$ on $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ given by $[t \mapsto f_1(t-f_2)]^*\mathcal{G}$, i.e., by

$$t \mapsto H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1(t-f_2))})).$$

Its geometric monodromy group G_{geom} contains SL.

proof The argument is quite similar to the one given for Theorem 5.4.9.We will indicate the modifications which must be made.

Put $r := \mathrm{rank}(\mathcal{F})$, $m := \#(\mathrm{Sing}(\mathcal{F})\cap(C-D_2))$, $\mathcal{F}_1 := f_{2*}(\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$. We have seen in Proposition 5.3.7 that $\mathcal{H}$ is the restriction to

$$\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$$

of the middle additive convolution of $\mathcal{F}_1$ and $\mathcal{L}_\chi$.

We have seen above (end of the proof of 5.3.6) that by the Irreducible Induction Criterion 3.3.1, $\mathcal{F}_1$ is an irreducible middle extension sheaf. It is tame at ∞, because $\mathcal{F}$ is tame at all the poles of f_2, and the poles of f_2 all have odd order.

We know that the function f_2 has

$$g-1 + \Sigma(1+c_i)/2 \geq (d_2+1)/2 - 1 \geq (6g+4)/2 - 1 \geq 1$$

critical points, and as many critical values. Over each critical value α of f_2, $\mathcal{F}_1$ is lisse, so the local monodromy of $\mathcal{F}_1$ at α is quadratic of drop r, with scale a character ρ_α of $I(\alpha)$ of order 2 and Swan conductor 1:

$$\mathcal{F}_1(\alpha)/\mathcal{F}_1(\alpha)^{I(\alpha)} \cong r \text{ copies of } \rho_\alpha.$$

Over the m images $\delta = f_2(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})\cap(C-D_2)$, f_2 is finite etale, and β is the unique point of $\mathrm{Sing}(\mathcal{F})\cap(C-D_2)$ in the fibre, so the local monodromy of $\mathcal{F}_1$ at δ has drop $\leq$ r. More precisely, we have

$$\mathcal{F}_1(\delta)/\mathcal{F}_1(\delta)^{I(\delta)} \cong \mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)},$$

where we use f_2 to identify $I(\delta)$ with $I(\beta)$.

Over each of the d_1 images $\gamma = f_2(\zeta)$ of the zeroes of f_1, f_2 is finite etale, ζ is the only zero of f_1 in its f_2–fibre, and $\mathcal{F}$ is lisse. Thus $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$ is lisse at all but the point ζ in the fibre $f_2^{-1}(\gamma)$. At ζ the local monodomy of $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$ is quadratic of drop r, with scale the character $\mathcal{L}_{\chi(\text{uniformizer at } \zeta)}$ of $I(\zeta)$. Thus the local monodromy of $\mathcal{F}_1$ at γ is quadratic of drop r, with scale the character $\mathcal{L}_{\chi(x-\gamma)}$ of $I(\gamma)$.

At all other points of $\mathbb{A}^1$, i.e., on $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$, $\mathcal{F}_1$ is lisse. As $\mathcal{F}$ is everywhere tame on C, $\mathcal{F}_1$ is tame outside the critical values of f_2. Now form $\mathcal{H}$, the middle additive convolution of $\mathcal{F}_1$ with $\mathcal{L}_\chi$. Thus (by 4.1.10, 2d), 1b), and 1c)) $\mathcal{H}$ is tame at ∞, it is tame outside the critical values of f_2, and it is lisse on $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$. Its rank is given by (5.2.1, part 5))

$$\begin{aligned}\mathrm{rank}(\mathcal{H}) = (2g-2 + d)r \\ &+ \Sigma_{P_i \text{ in } D^{red}} \mathrm{Swan}_{P_i}(\mathcal{F}) + \Sigma_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{finite}} \mathrm{Swan}_s(\mathcal{F}) \\ &+ \Sigma_{s \text{ in } \mathrm{Sing}(\mathcal{F})_{finite}} \mathrm{drop}_s(\mathcal{F}) \\ &+ \Sigma_{P_i \text{ in } D^{red}} \mathrm{drop}_{P_i}(\mathcal{F}(P_i)\otimes(\mathcal{L}_\chi a_i)(\infty, P_i)),\end{aligned}$$

where we have written $\mathrm{Sing}(\mathcal{F})_{finite}$ for $\mathrm{Sing}(\mathcal{F})\cap(C-D)$.
In particular, we have the inequality (5.2.1, part 6))

$$\mathrm{rank}(\mathcal{H}) \geq (2g-2 + d)r.$$

The local monodromy of $\mathcal{H}$ at the m images $\delta = f_2(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})\cap(C-D_2)$ is tame and has drop $\leq$ r, by 4.1.10 parts 1b) and 1c).

The local monodromy of $\mathcal{H}$ at each critical value α of f_2 is quadratic of drop r, with scale a character $MC_\chi\mathrm{loc}(\alpha)(\rho_\alpha)$ whose order, twice the order of χ by 4.2.2, is ≥ 6. Thus

$$\mathcal{H}(\alpha)/\mathcal{H}(\alpha)^{I(\alpha)} \cong \text{r copies of a character of order} \geq 6.$$

Over each of the d_1 images $\gamma = f_2(\zeta)$ of the zeroes of f_1, the local monodromy of $\mathcal{H}$ at γ is quadratic of drop r, with scale the character $\mathcal{L}_{\chi^2(x-\gamma)}$ of $I(\gamma)$, whose order, that of χ, is ≥ 3.

With the exception of at most m points of $\mathbb{A}^1$, namely the images by f_2 of points in $\mathrm{Sing}(\mathcal{F})\cap(C-D_2)$, the local monodromy of $\mathcal{H}$ is quadratic of drop r, with scale a character not of order 2. The conclusion follows from Theorem 1.5.1 (and Theorem 1.7.1, if r=1), applied to (r, m, $\mathcal{H}$), with $S - S_0$ the critical values of f together with the d_1 images $\gamma = f_2(\zeta)$ of the zeroes of f_1, and S_0 the m images $\delta = f(\beta)$ of points β in $\mathrm{Sing}(\mathcal{F})\cap(C-D_2)$. QED

5.7 Theorems of big monodromy for $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ on *Fct*(C, d, D, $\mathrm{Sing}(\mathcal{F})_{finite}$) in characteristic 2

Theorem 5.7.1 Let k be an algebraically closed field of characteristic 2. Fix a prime number ℓ which is invertible in k. Fix a nontrivial character χ of finite odd order $n \geq 3$. Let C/k be a proper smooth connected curve of genus g. Fix an irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on C. Let $D = \Sigma a_i P_i$ be an effective divisor of degree d on C. Suppose that either

1a) $d \geq 6g+3$, all a_i are odd, $\mathrm{Sing}(\mathcal{F})\cap(C-D)$ is nonempty, and the following inequalities hold:

if $\mathrm{rank}(\mathcal{F}) = 1$, $2g - 2 + d \geq \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D)), 4\mathrm{rank}(\mathcal{F}))$,
if $\mathrm{rank}(\mathcal{F}) \geq 2$, $2g - 2 + d \geq \mathrm{Max}(2\#(\mathrm{Sing}(\mathcal{F})\cap(C-D)), 72\mathrm{rank}(\mathcal{F}))$,

or
1b) $d \geq 12g+7$, and the following inequalities hold:

if $\mathrm{rank}(\mathcal{F}) = 1$, $2g - 2 + d > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}), 4\mathrm{rank}(\mathcal{F}))$.
if $\mathrm{rank}(\mathcal{F}) \geq 2$, $2g - 2 + d > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}), 72\mathrm{rank}(\mathcal{F}))$.

Suppose further that
2) $\mathcal{F}$ is everywhere tame.

Then for the lisse sheaf $\mathcal{G}$ on $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ given by

$$f \mapsto H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}),$$

G_{geom} contains SL.

proof This follows from Theorems 5.6.1 and 5.6.2 above in exactly the same way that Theorem 5.5.1 followed from Theorems 5.4.1 and 5.4.9. QED

Chapter 6: Dependence on Parameters

6.0 A lemma on relative Cartier divisors

(6.0.1). The following lemma is standard. We include it for ease of reference.

Lemma 6.0.2 Let T be an arbitrary scheme, X/T a proper smooth T−scheme with geometrically connected fibres everywhere of dimension N, $\mathcal{L}$ an invertible $\mathcal{O}_X$−module, and L in $H^0(X, \mathcal{L})$ a global section. Suppose L is nonzero on each geometric fibre of X/T, i.e., for every geometric point t of T, the image L_t of L in $H^0(X_t, \mathcal{L}_t)$ is nonzero. Then the locus "L = 0 as section of $\mathcal{L}$", call it Z, is a Cartier divisor in X, which is flat over T.

proof The question is Zariski local on T, which we may assume affine, say T = Spec(R). All the data (X/R, Z/R, L) is of finite presentation over R, so we may reduce to the case where R is noetherian, then to the case where R is noetherian local, then to the case where R is complete noetherian local, and finally to the case where R is complete noetherian local with algebraically closed residue field k.

It suffices to show that, over any such R, the sheaf map

$$\mathcal{L}^{-1} \xrightarrow{\times L} \mathcal{O}_X$$

is injective on X. Indeed, for any ideal I in R, R/I is again complete noetherian local with algebraically closed residue field, so after the base change R → R/I we will again have the injectivity of

$$\mathcal{L}^{-1}/I\mathcal{L}^{-1} \xrightarrow{\times L} \mathcal{O}_X/I\mathcal{O}_X$$

This means precisely that the short exact sequence

$$0 \to \mathcal{L}^{-1} \xrightarrow{\times L} \mathcal{O}_X \to \mathcal{O}_X/f\mathcal{L}^{-1} = \mathcal{O}_Z \to 0$$

remains exact after any base change R → R/I. Because $\mathcal{L}^{-1}$ and $\mathcal{O}_X$ are flat over R, the Tor sequence gives a four−term exact sequence

$$0 \to \mathrm{Tor}_1{}^R(\mathcal{O}_Z, R/I) \to \mathcal{L}^{-1}/I\mathcal{L}^{-1} \to \mathcal{O}_X/I\mathcal{O}_X \to \mathcal{O}_Z/I\mathcal{O}_Z \to 0.$$

Therefore $\mathrm{Tor}_1{}^R(\mathcal{O}_Z, R/I) = 0$ for any ideal I in R, i.e., $\mathcal{O}_Z$ is flat over R, as required.

To show that multiplication by $L : \mathcal{L}^{-1} \to \mathcal{O}_X$ is injective on X, we argue as follows. If not, there is some closed point x in X over whose complete local ring $\mathcal{O}_{X,x}{}^\wedge$ the map

$$L : \mathcal{L}^{-1} \otimes_{\mathcal{O}_X} \mathcal{O}_{X,x}{}^\wedge \to \mathcal{O}_{X,x}{}^\wedge$$

is not injective. If we pick a basis e of the source, which is a free, rank one $\mathcal{O}_{X,x}{}^\wedge$−module, then

Le is an element of $\mathcal{O}_{X,x}{}^\wedge$ which is nonzero in $\mathcal{O}_{X,x}{}^\wedge/\mathcal{M}_R\mathcal{O}_{X,x}{}^\wedge$. We must show that Le is not a zero divisor in $\mathcal{O}_{X,x}{}^\wedge$. The closed point x in X lies over the closed point of Spec(R), so x has residue field k. Because X/T is smooth of relative dimension N, there exists $\tilde{x}$ in X(R) which lifts x in X(k), and we have an isomorphism of local rings

$$\mathcal{O}_{X,x}{}^\wedge \cong R[[X_1, ..., X_N]].$$

Our element Le in $R[[X_1, ..., X_N]]$, say $Le \approx \sum_w r_w X^w$, reduces mod $\mathcal{M}_R$ to a nonzero element of $k[[X_1, ..., X_N]]$. We claim that any such element of $R[[X_1, ..., X_N]]$ is not a zero divisor.

This is an elementary application of the Weierstrass preparation theorem. At least one of its coefficients r_w is a unit in R. The minimum $|w|$ such that r_w is a unit in R is the "Weierstrass degree" of $\sum_w r_w X^w$, call it n. After a suitable linear change of variables, we may assume the monomial $(X_N)^n$ occurs with coefficient 1. Now view $R[[X_1, ..., X_N]]$ as $R_{N-1}[[X_N]]$, with R_{N-1} the power series ring $R[[X_1, ..., X_{N-1}]]$. By the Weierstrass Preparation Theorem, the element Le is the product of a unit with a Weierstrass polynomial in X_N of degree n,

$$(X_N)^n + \sum_{i \le n-1} m_i (X_N)^i$$

with all m_i in the maximal ideal of R_{N-1}. But no Weierstrass polynomial in X_N is a zero divisor in $R_{N-1}[[X_N]]$. Indeed, suppose for some g in $R_{N-1}[[X_N]]$ we have

$$((X_N)^n + \sum_i m_i (X_N)^i) g = 0,$$

then

$$(X_N)^n g = -(\sum_i m_i (X_N)^i) g.$$

Suppose we have already established that g has all coefficients in the k'th power of the maximal ideal of R_{N-1}. Then the equation above shows that $(X_N)^n g$, and hence g itself, has all coefficients in the k+1'st power. Proceeding in this way, we conclude that all coefficients of g lie in $\cap_k (\mathcal{M}_{R_{N-1}})^k = \{0\}$. QED

6.1 The situation with curves

(6.1.1) We fix an arbitrary scheme T, which will play the role of a parameter space in what follows. We fix an integer $g \ge 0$, and a relative curve C/T of genus g. More precisely, we fix

(6.1.1.1) $$\pi : C \to T$$

a proper smooth morphism whose fibres are geometrically connected curves of genus g. We suppose given an integer $d \ge 2g-1$ and an effective Cartier divisor D in C which is finite and flat over T of degree d.

Lemma 6.1.2 Let T be a scheme, $g \ge 0$ an integer, and

$$\pi : C \to T,$$

a proper smooth morphism whose fibres are geometrically connected curves of genus g. Suppose given an integer $d \geq 1$ and an effective Cartier divisor D in C which is finite and flat over T of degree d. Suppose we are given a global section f of $H^0(C, I^{-1}(D))$ which is nonzero on each geometric fibre of C/T. Then the locus "f=0 as section of $I^{-1}(D)$", call it Z, is an effective Cartier divisor in C, finite and flat over T of rank d.

proof We already know that Z/T is a relative Cartier divisor in C/T, flat over T. Because Z is closed in C, Z is proper over T. Then Z/T is finite, because it has finite fibres. Thus Z/T is finite and flat. One sees that it is finite and flat of degree d by looking at fibres. QED

Lemma 6.1.3 Hypotheses as in Lemma 6.1.2, suppose in addition that $d \geq 2g - 1$. Consider the functor on T–schemes Y/T given by

$$Y/T \mapsto \text{the set of global sections of } H^0(C_Y, I^{-1}(D)_Y) \text{ which are nonzero on each}$$

geometric fibre of C_Y/Y.

This functor is represented by a T–scheme $L(D)_{nonzero}/T$, namely the complement of the zero section in the total space of the vector bundle on T of rank d+1–g given by $\pi_*(I^{-1}(D))$.

proof The only point is that because $d > 2g-2$, $\pi_*(I^{-1}(D))$ on T is a locally free $\mathcal{O}_T$–module whose formation commutes with arbitrary change of base on T. QED

Definition 6.1.4 Hypotheses as in Lemma 6.1.2 above, a global section f of $H^0(C, I^{-1}(D))$ is said to have d distinct zeroes if it is nonzero on each geometric fibre of C/T and if Z, the locus "f=0 as section of $I^{-1}(D)$", is finite etale over Y.

Lemma 6.1.5 Hypotheses as in Lemma 6.1.2, suppose in addition that $d \geq 2g - 1$. Consider the functor on T–schemes Y/T given by

$$Y/T \mapsto \text{the set of global sections of } H^0(C_Y, I^{-1}(D)_Y) \text{ which have d distinct zeroes.}$$

This functor is represented by a T–scheme $L(D)_{d\ dist\ zeroes}/T$, which is an open set in $L(D)_{nonzero}/T$.

proof If we make the base change from T to $Y := L(D)_{nonzero}/T$, we acquire the universal global section f_{univ} which is nonzero on geometric fibres. Over this base space Y, we have the finite flat scheme Z/Y. Its structure sheaf $\mathcal{O}_Z$ is an $\mathcal{O}_Y$–algebra which is a locally free $\mathcal{O}_Y$–module of rank d. Then $L(D)_{d\ dist\ zeroes}/T$ is the open subscheme of Y over which Z/Y is finite etale. Locally on Y, if we pick an $\mathcal{O}$ basis of $\mathcal{O}_Z$, say $e_1, ..., e_d$, $L(D)_{d\ dist\ zeroes}/T$ is the open set where the discriminant

$$\det_{d\times d}(\mathrm{Trace}_{\mathcal{O}_Y}(e_ie_j))$$

is invertible. QED

Definition 6.1.6 Hypotheses as in Lemma 6.1.2 above, we say a global section f of $H^0(C, I^{-1}(D))$ is invertible near D, or has exact divisor of poles D, if the following condition is satisfied. Multiplication by f defines an $\mathcal{O}_C$–linear map

$$\mathcal{O}_C/I(D) \xrightarrow{\times f} I^{-1}(D)/\mathcal{O}_C.$$

Taking π_*, we get an $\mathcal{O}_T$–linear map "f|D"

$$f|D : \pi_*(\mathcal{O}_C/I(D)) \to \pi_*(I^{-1}(D)/\mathcal{O}_C)$$

between locally free $\mathcal{O}_T$–modules of the same rank d. We require that f|D be an isomorphism. [If locally on T we take $\mathcal{O}_T$–bases of source and target, we can calculate the determinant of f|D. Locally on T, this determinant is well defined in $\mathcal{O}_T$, up to multiplication by an invertible section of $\mathcal{O}_T$.We require that this determinant be everywhere invertible on T.]

Lemma 6.1.7 Hypotheses as in Lemma 6.1.2, suppose in addition that $d \geq 2g - 1$. Consider the functor on T–schemes Y/T given by

Y/T $\mapsto$ the set of global sections of $H^0(C_Y, I^{-1}(D)_Y)$ which are invertible near D_Y.

This functor is represented by a T–scheme $L(D)_{\text{inv near D}}/T$. Locally on T, $L(D)_{\text{inv near D}}/T$ is a principal open set in $L(D)_{\text{nonzero}}/T$.

proof If we make the base change from T to Y := $L(D)_{\text{nonzero}}/T$, we acquire the universal global section f_{univ} which is nonzero on geometric fibres. Over this base space Y, we have the map $f_{\text{univ}}|D_Y$

$$f_{\text{univ}}|D_Y : \pi_{Y*}(\mathcal{O}_{C_Y}/I(D_Y)) \to \pi_{Y*}(I^{-1}(D_Y)/\mathcal{O}_{C_Y})$$

of locally free $\mathcal{O}_Y$–modules of rank d. Our functor is represented by the open set of Y where the "determinant" of $f_{\text{univ}}|D_Y$ is invertible. QED

Definition 6.1.8 Hypotheses as in Lemma 6.1.2, suppose we are given in addition an integer $s \geq 0$ and an effective Cartier divisor S in C/T, which is finite and flat over T of degree s (with the convention that S is empty if s = 0), and which is scheme–theoretically disjoint from D. A global section f of $H^0(C, I^{-1}(D))$ is said to be invertible near S if the following conditions hold. If s = 0, we require only that f be nonzero on each geometric fibre of C/T. If $s \geq 1$, multiplication by f defines an $\mathcal{O}_C$–linear endomorphism of $\mathcal{O}_S := \mathcal{O}_C/I(S)$. Taking π_*, we get an $\mathcal{O}_T$–linear endomorphism "f|S"

$$f|S : \pi_*(\mathcal{O}_C/I(S)) \to \pi_*(\mathcal{O}_C/I(S))$$

of locally free $\mathcal{O}_T$–modules of the same rank s. We require that f|S be an isomorphism. Here we have a true endomorphism, so we can speak of det(f|S) as a global section of $\mathcal{O}_T$. We require that this determinant be an invertible global section of $\mathcal{O}_T$.

Lemma 6.1.9 Hypotheses as in Lemma 6.1.2, suppose we are given in addition an integer $s \geq 0$ and an effective Cartier divisor S in C/T, which is finite and flat over T of degree s (with the convention that S is empty if $s = 0$), and which is scheme–theoretically disjoint from D. Suppose $d \geq 2g - 1$. Consider the functor on T–schemes Y/T given by

Y/T $\mapsto$ the set of global sections of $H^0(C_Y, I^{-1}(D)_Y)$ which are invertible near D_Y **and** invertible near S_Y.

This functor is represented by a T–scheme $L(D)_{\text{inv near D and S}}/T$, which is a principal open set in $L(D)_{\text{inv near D}}/T$ for $s \geq 1$, and which is equal to $L(D)_{\text{inv near D}}/T$ for $s = 0$.

proof If $s = 0$, there is nothing to prove. If $s \geq 1$, make the base change from T to $Y := L(D)_{\text{inv near D}}/T$. We acquire the universal global section f_{univ} which is invertible near D. Over this base space Y, our functor is represented by the open set of Y where $\det(f_{\text{univ}}|S_Y)$ is invertible. QED

Lemma 6.1.10 Hypotheses as in Lemma 6.1.2, suppose we are given in addition an integer $s \geq 0$ and an effective Cartier divisor S in C/T, which is finite and flat over T of degree s (with the convention that S is empty if $s = 0$), and which is scheme–theoretically disjoint from D. Suppose that $d \geq 2g - 1$. Consider the functor on T–schemes Y/T given by

Y/T $\mapsto$ the set of global sections f of $H^0(C_Y, I^{-1}(D)_Y)$ which are invertible near D_Y **and** invertible near S_Y, and which have d distinct zeroes.

This functor is represented by a T–scheme

$$\mathcal{F}ct(C, d, D, S)$$

which is open in both $L(D)_{\text{inv near D and S}}$ and in $L(D)_{\text{d dist zeroes}}$.

proof Indeed, the functor $\mathcal{F}ct(C/T, d, D, S)$ is represented by the fibre product over $L(D)_{\text{nonzero}}$ of the open subschemes

$$L(D)_{\text{inv near D and S}} \times_{L(D)_{\text{nonzero}}} L(D)_{\text{d dist zeroes}}.\ \text{QED}$$

Remark 6.1.11 When T is the spec of an algebraically closed field, the set of k–valued points of the k–scheme *Fct*(C, d, D, S) is precisely the space Fct(C, d, D, S). The possibility of taking T to be the spec of a finite field k will be absolutely essential in the chapters which follow.

6.2 Construction of the twist sheaf $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ with parameters

(6.2.1) We fix a prime number ℓ, and a normal and connected $\mathbb{Z}[1/\ell]$–scheme T. We assume further that T is a "good scheme" in the sense of [Ka–RLS, 4.0], i.e., that T admits a map of finite type to a scheme which is regular of dimension ≤ 1. We fix an integer $g \geq 0$, and a curve C/T of genus g, i.e., we fix

(6.2.1.1) $$\pi : C \to T,$$

a proper smooth morphism whose fibres are geometrically connected curves of genus g.

(6.2.2) We suppose given an integer $d_0 \geq 1$ and an effective Cartier divisor D_0 in C which is finite etale over T of degree d_0. We further suppose given an integer $d \geq 2g+1$ and an effective Cartier divisor D in C which is finite and flat over T of degree d, such that

(6.2.2.1) $$D^{\mathrm{red}} = (D_0)^{\mathrm{red}}.$$

[Thus etale locally on T, D_0 is a disjoint union of sections, $D_0 = \amalg_i P_i$, and the divisor D is $\Sigma a_i P_i$ for some choice of strictly positive integers a_i with $\Sigma_i a_i = d$.]

(6.2.3) We also suppose given an integer $s \geq 0$ and an effective Cartier divisor S in C – D which is finite etale over T of degree s, with the convention that if s = 0 then S is empty. We may also view S as an effective Cartier divisor in C which is finite etale over T of degree s, and which is disjoint from D_0. [Thus etale locally on T, S is a disjoint union of sections, $S = \amalg_j Q_j$, D_0 is a disjoint union of sections, $D_0 = \amalg_i P_i$, the divisor D is $\Sigma a_i P_i$, and for all i and j, P_i and Q_j are disjoint.]

(6.2.4) Our last data is an integer $r \geq 1$ and a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ of rank r on C – D – S, about which we make the following two hypotheses:

(6.2.4.1) For each geometric point t of T, the lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}_t$ or rank r on $C_t - D_t - S_t$ is irreducible.

(6.2.4.2) For variable geometric points t of T, the compact Euler characteristic $\chi_c(C_t - D_t - S_t, \mathcal{F}_t)$ is a constant function of t.

(6.2.5) Notice that all of the conditions we have imposed are stable under arbitrary change of base on T.

Remark 6.2.6 If, for each geometric point t in T, the lisse sheaf $\mathcal{F}_t$ on the open curve $C_t - D_t - S_t$ is everywhere tame, then condition 6.2.4.2 holds trivially, for then

$$\chi_c(C_t - D_t - S_t, \mathcal{F}_t) = r\chi_c(C_t - D_t - S_t) = r(2 - 2g - d_0 - s).$$

If the generic point of our normal connected scheme T is (the spectrum of) a field of characteristic zero, this tameness is automatic, cf. [Ka–SE, 4.7.1].

Remark 6.2.7 To understand better condition 6.2.4.2 in a less trivial case, suppose in addition that the divisors D_0 and S are disjoint unions of sections of C/T, say $D_0 = \coprod_i P_i$ and $S = \coprod_j Q_j$, and that the divisor D is $\sum a_i P_i$. By the Euler–Poincaré formula, we have

$$\begin{aligned}&\chi_c(C_t - D_t - S_t, \mathcal{F}_t)\\ &= r\chi_c(C_t - D_t - S_t) - \Sigma_i \, \mathrm{Swan}_{P_i(t)}(\mathcal{F}_t) - \Sigma_j \, \mathrm{Swan}_{Q_j(t)}(\mathcal{F}_t)\\ &= r(2 - 2g - d_0 - s) - \Sigma_i \, \mathrm{Swan}_{P_i(t)}(\mathcal{F}_t) - \Sigma_j \, \mathrm{Swan}_{Q_j(t)}(\mathcal{F}_t).\end{aligned}$$

So condition 6.2.4.2 certainly holds if each of the Swan terms $\mathrm{Swan}_{P_i(t)}(\mathcal{F}_t)$ and $\mathrm{Swan}_{Q_j(t)}(\mathcal{F}_t)$ is a constant function of t. By Deligne's semicontinuity theorem [Lau–SC, 2.1.1], each of these Swan terms separately is constructible and lower semicontinuous in t. Therefore 6.2.4.2 holds if and only if each Swan term is itself a constant function of t.

(6.2.8) Now choose an integer n invertible on T, and suppose T is given a structure of $\mathbb{Z}[1/n, \zeta_n]$–scheme. (Here we write $\mathbb{Z}[1/n, \zeta_n]$ for the ring $\mathbb{Z}[1/n, X]/(\Phi_n(X))$, where $\Phi_n(X)$ denotes the n'th cyclotomic polynomial.) Given a character

(6.2.8.1) $$\chi : \mu_n(\mathbb{Z}[1/n, \xi_n]) \to (\overline{\mathbb{Q}}_\ell)^\times$$

of order n, we get a lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf on $\mathbb{G}_m/\mathbb{Z}[1/n, \zeta_n]$ by pushing out by χ the Kummer torsor

$$\begin{aligned}[-n] : \mathbb{G}_m &\to \mathbb{G}_m,\\ x &\mapsto x^{-n},\end{aligned}$$

whose structural group is $\mu_n(\mathbb{Z}[1/n, \xi_n])$. By pullback, we get $\mathcal{L}_\chi$ on $\mathbb{G}_m/T$.

From the data (C/T, D, S) we construct the space

$$X := \mathit{Fct}(C, d, D, S)/T.$$

On $C_X := C \times_T X$, we have the universal section f of $I^{-1}(D_X)$, its zero locus Z/X, and the open curve

$$C_X - D_X - S_X - Z.$$

If we think of f as a section of the structural sheaf of $C_X - D_X$, then we may view

$$C_X - D_X - S_X - Z$$

as being

$$(C_X - D_X - S_X)[1/f].$$

Then f is an invertible function on $C_X - D_X - S_X - Z$, so we may form the lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{L}_{\chi(f)} := f^*\mathcal{L}_\chi$.

(6.2.9) We denote by

$$p : C_X - D_X - S_X - Z \to X$$

the structural morphism, just as in 5.2.1 (but p was denoted π there).

Proposition 6.2.10 Given data (C/T, D, S, ℓ, r, $\mathcal{F}, \chi$) satisfying all the hypotheses made above in 6.2.1–4 and 6.2.8–9, we have the following results.
1) The sheaves $R^ip_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on X vanish for $i\neq1$, and $R^1p_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is lisse.
2) The sheaves $R^ip_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on X vanish for $i\neq1$, and $R^1p_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is lisse, and of formation compatible with arbitrary change of base.
3) The image $\mathcal{G}$ of the natural "forget supports" map

$$R^1p_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) \to R^1p_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$$

is lisse, of formation compatible with arbitrary change of base on X. In particular, the formation of $\mathcal{G}$ commutes with arbitrary base change on T. Thus when we base change to a geometric point of T, i.e., to a point of T with values in the spec of an algebraically closed field k, we recover the construction of 5.2.1.
4) If, for some integer w, the lisse sheaf $\mathcal{F}$ on C – D – S carries an orthogonal (respectively symplectic) autoduality toward $\overline{\mathbb{Q}}_\ell(-w)$,

$$< . > : \mathcal{F}\times\mathcal{F} \to \overline{\mathbb{Q}}_\ell(-w),$$

and χ has order two, then the Poincaré duality pairing on X,

$$R^1p_!(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}) \times R^1p_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$$
$$\to R^2p_!(\mathcal{F}\otimes\mathcal{F}) \to R^2p_!(\overline{\mathbb{Q}}_\ell(-w)) \cong \overline{\mathbb{Q}}_\ell(-w-1),$$

deduced from cup product and < . >, induces on $\mathcal{G}$ a symplectic (respectively orthogonal) autoduality toward $\overline{\mathbb{Q}}_\ell(-w-1)$ on X,

$$< , > : \mathcal{G}\times\mathcal{G} \to \overline{\mathbb{Q}}_\ell(-w-1).$$

proof Simply repeat the proof of 5.2.1. QED

Chapter 7: Diophantine Applications over a Finite Field

7.0 The general setup over a finite field: relation of the sheaf $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$ to L–functions of twists

(7.0.1) In this section, we work over a **finite** field k, of cardinality q and characteristic p. We fix a proper, smooth, geometrically connected curve C/k of genus g, an effective divisor D on C of degree $d \geq 2g+1$, a prime number ℓ invertible in k, an integer $r \geq 1$, and a geometrically irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on C of generic rank r. We denote by $\mathrm{Sing}(\mathcal{F}) \subset C$ the finite set of closed points of C at which $\mathcal{F}$ is not lisse, and by $\mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ the intersection $\mathrm{Sing}(\mathcal{F}) \cap (C-D)$. The space

(7.0.1.1) $$X := \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$$

has a natural structure of scheme over k, cf. Proposition 6.1.10. For any extension field E/k, the E–valued points X(E) consist of those functions f in $H^0(C \otimes_k E, I^{-1}(D))$ whose divisor of zeroes $f^{-1}(0)$ is both disjoint from $D \cup \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}$ and finite etale of degree d over E.

(7.0.2) We also fix a nontrivial $\overline{\mathbb{Q}}_\ell$–valued multiplicative character

(7.0.2.1) $$\chi: k^\times \to \overline{\mathbb{Q}}_\ell^\times,$$

and denote by $\mathcal{L}_\chi$ the corresponding Kummer sheaf on $\mathbb{G}_m/k$.

(7.0.3) The construction 5.2.1, carried out over the finite field k instead of over $\overline{k}$, provides us with a lisse $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$$

on $X := \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$, cf. Proposition 6.2.10.

(7.0.4) The fundamental diophantine property of $\mathcal{G}$ is this. Given any finite extension field E/k inside $\overline{k}$, and any f in X(E), the stalk $\mathcal{G}_f$ of $\mathcal{G}$ at (the geometric point "f as $\overline{k}$–valued point" lying over) f is the cohomology group

$$\mathcal{G}_f = H^1(C \otimes_k \overline{k}, j_*(\mathcal{F} \otimes \mathcal{L}_{\chi(f)})),$$

and the action of $\mathrm{Frob}_{E,f}$ on $\mathcal{G}_f$ is the action of Frob_E on this cohomology group. Thus we have

$$\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}) = \det(1 - T\mathrm{Frob}_E \mid H^1(C \otimes_k \overline{k}, j_*(\mathcal{F} \otimes \mathcal{L}_{\chi(f)}))).$$

Remark 7.0.5 By Chebotarev, any lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{H}$ on X is determined up to semisimplification by all its local characteristic polynomials of Frobenius $\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{H})$. Applying this fact to $\mathcal{G}$, and remembering that $\mathcal{G}$ is irreducible, we see that $\mathcal{G}$ is in fact determined up to isomorphism by its fundamental diophantine property.

(7.0.6) We can also think of $\mathcal{G}$ as the sheaf whose local characteristic polynomials at E–valued points f in X(E), (7.0.6.1) $\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G})$,

are the global L–functions of $C\otimes_k E$ with coefficients in $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$. Indeed, the sheaf $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ on $C\otimes_k E$ is a geometrically irreducible middle extension, which is not geometrically constant (because f has simple zeroes at points where $\mathcal{F}$ is lisse). Therefore we have

(7.0.6.2) $H^i(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})) = 0$ for $i \neq 1$.

The L–function of $C\otimes_k E$ with coefficients in $j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})$ is, by the Lefschetz Trace Formula, given by the alternating product

(7.0.6.3)
$$L(C\otimes_k E, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))(T)$$
$$= \prod_{i=0,1,2}\det(1 - T\mathrm{Frob}_E \mid H^i(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})))^{(-1)^{i+1}}.$$

In view of the above vanishing (7.0.6.2) of H^i for $i \neq 1$, we have

(7.0.6.4)
$$L(C\otimes_k E, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))(T)$$
$$= \det(1 - T\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})))$$
$$= \det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}).$$

(7.0.7) Put

(7.0.7.1) $N := \mathrm{rank}(\mathcal{G})$.

(7.0.8) We fix an embedding $\iota: \overline{\mathbb{Q}}_\ell \to \mathbb{C}$. We further suppose that $\mathcal{F}$ is ι–pure, of integer weight denoted w. This means that for every finite extension E of k, every E–valued point x of C at which $\mathcal{F}$ is lisse, and every eigenvalue λ of $\mathrm{Frob}_{E,x}$ on $\mathcal{F}$, we have

$$|\iota(\lambda)| = (\#E)^{w/2}.$$

[Recall that $\mathcal{F}$ is said to be pure of weight w if it is ι–pure of weight w for every choice of ι.] Because $\mathcal{F}$ is ι–pure of weight w, $\mathcal{G}$ is ι–pure of weight w+1, thanks to Deligne [De–WeII, 3.2.3].

(7.0.9) We also fix a choice α_k of $(\#k)^{-1/2}$ in $\overline{\mathbb{Q}}_\ell$, which may or may not map by ι to the positive square root. This choice allows us to perform Tate twists by half–integers. In the notation of [Ka–Sar, RMFEM, 9.0.11], $\mathcal{F}(n/2)$ is $\mathcal{F}\otimes\beta^{\deg}$, for $\beta := (\alpha_k)^n$. Thus $\mathcal{F}(w/2)$ and $\mathcal{G}((w+1)/2)$ are both ι–pure of weight zero.

7.1 Applications to equidistribution

(7.1.1) Suppose we are given data (C/k, D, ℓ, r, $\mathcal{F}$, χ, ι, w) as in the previous section 7.0. We wish to apply Deligne's general equidistribution theorem [De–WeII, 3.5.3], cf. also [Ka–GKM, 3.6] and [Ka–Sar, RMFEM, 9.2.6], to $\mathcal{G}$. For this, we need to know the group G_{geom} for $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\mathcal{F})$. To this end, we suppose that after extension of scalars from k to $\bar{k}$, our data (C/k, D, ℓ, r, $\mathcal{F}$, χ) satisfies all the hypotheses of Theorem 5.5.1, if char(k) is odd, or of Theorem 5.7.1, if char(k) is two. Thus G_{geom} for $\mathcal{G}$ is either Sp(N) or SO(N) or O(N) or a group containing SL(N). We now discuss each of these cases separately, in order of increasing complexity.

7.2 The SL case

(7.2.1) Let us first examine in greater detail the case when G_{geom} contains SL(N) (and the hypotheses of section 7.0 are in force). Because $\mathcal{G}$ lives over a finite field k and is irreducible, we know [De–WeII, 1.3.9] that G_{geom} is a semisimple group. But the only semisimple groups between SL(N) and GL(N) are the groups

(7.2.1.1) $$GL_\nu(N) := \{A \text{ in } GL(N) \mid \det(A)^\nu = 1\}$$

for $\nu \geq 1$ an integer. Therefore for some integer $\nu \geq 1$ we have

(7.2.1.2) $$G_{geom} = GL_\nu(N).$$

(7.2.2) Suppose that the parameter space X admits a k–rational point f. Then if we twist $\mathcal{G}$ by an N'th root β of $1/\det(Frob_{k,f} \mid \mathcal{G}_f)$, the resulting lisse sheaf $\mathcal{G}\otimes\beta^{deg}$ is ι–pure of weight zero, and all its Frobenii land in G_{geom}. We should remark here that the quantity

(7.2.2.1) $$(-1)^N\det(Frob_{k,f} \mid \mathcal{G}_f) = \det(-Frob_{k,f} \mid \mathcal{G}_f)$$
$$= \det(-Frob_k \mid H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})))$$

is the constant in the functional equation for the L–function

(7.2.2.2) $$L(C\otimes_k E, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}))(T).$$

As such, it is a product, over the closed points of C, of local constants, cf. [De–Const] and [Lau–TFC]. At least in favorable cases, these local constants are eminently computable, cf. 7.9.5 and 8.9.2. In this sense, the recipe in 7.2.2 above for β is an "explicit" one.

(7.2.3) We take

(7.2.3.1) $$K := U_\nu(N) := \{A \text{ in } U(N) \mid \det(A)^\nu = 1\},$$

a maximal compact subgroup of $G_{geom}(\mathbb{C})$. For each finite extension E/k inside $\bar{k}$, and each f in X(E), we denote by $\theta(E, f)$ the Frobenius conjugacy class in K attached to $\mathcal{G}\otimes\beta^{deg}$ at the E–valued point f of X. Thus

(7.2.3.2) $$\det(1 - T\theta(E, f)) := \iota(\det(1 - TFrob_{E,f} \mid \mathcal{G}\otimes\beta^{deg}))$$
$$= \iota(\det(1 - T\beta^{deg(E/k)}Frob_{E,f} \mid H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})))).$$

(7.2.4) This equality 7.2.3.2 of characteristic polynomials determines $\theta(E, f)$ as a conjugacy class in K. By Deligne's general equidistribution theorem [De–WeII, 3.5.3], cf. also [Ka–GKM, 3.6] and [Ka–Sar, RMFEM, 9.2.6], as $\#E \to \infty$, the conjugacy classes

$$\{\theta(E, f)\}_{f \text{ in } X(E)}$$

become equidistributed for Haar measure in the space $K^\#$ of conjugacy classes in K.

(7.2.5) What happens if we do not assume that the parameter space X admits a k–rational point? We can still prove the existence of a β such that all Frobenii for $\mathcal{G}\otimes\beta^{deg}$ land in G_{geom}. Simply replace $\mathrm{Frob}_{k,f}$ by any element γ of $\pi_1(X)$ which maps onto Frob_k in $\pi_1(\mathrm{Spec}(k)) = \mathrm{Gal}(\bar{k}/k)$, and take for β an N'th root of $1/\det(\gamma \mid \mathcal{G})$. For any such β, $\mathcal{G}\otimes\beta^{\deg}$ is ι–pure of weight zero (because for an ι–pure lisse sheaf, its weight is equal to its determinental weight, cf. [De–WeII, 1.3.5].

(7.2.6) Here is a more concrete version of the above recipe for a suitable β. For each $n \geq 1$, denote by $k_n \subset \bar{k}$ the extension of k of degree n. For each $n >> 0$, X has a k_n–valued point, say f_n. Take β an N'th root of

(7.2.6.1) $$\det(\mathrm{Frob}_{k_n,f_n} \mid \mathcal{G})/\det(\mathrm{Frob}_{k_{n+1},f_{n+1}} \mid \mathcal{G}).$$

7.3 The Sp case

(7.3.1) Let us next consider the case in which $\mathcal{F}(w/2)$ is orthogonally self–dual on C/k, and χ has order 2 (and the hypotheses of section 7.0 are in force). Then, by Poincaré duality, $\mathcal{G}((w+1)/2)$ is symplectically self–dual on X. The field k must have char(k) $\neq 2$, simply because χ has order 2. By hypothesis, Theorem 5.5.1 holds, so $\mathcal{G}$ has $G_{geom} = Sp(N)$. Thus the lisse sheaf $\mathcal{G}((w+1)/2)$ is ι–pure of weight zero, and all its Frobenii land in G_{geom}. In this case we take

(7.3.1.1) $$K := USp(N),$$

a maximal compact subgroup of $G_{geom}(\mathbb{C})$. For each finite extension E/k inside $\bar{k}$, and each f in X(E), we denote by $\theta(E, f)$ the Frobenius conjugacy class in K attached to $\mathcal{G}((w+1)/2)$ at the E–valued point f of X. Thus

(7.3.1.2) $$\det(1 - T\theta(E, f)) := \iota(\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)))$$
$$= \iota(\det(1 - T\alpha_k^{\deg(E/k)(w+1)}\mathrm{Frob}_{E,f} \mid H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})))).$$

(7.3.2) This equality 7.3.1.2 of characteristic polynomials determines $\theta(E, f)$ as a conjugacy class in K. By Deligne's general equidistribution theorem [De–WeII, 3.5.3], cf. also [Ka–GKM, 3.6] and [Ka–Sar, RMFEM, 9.2.6], as $\#E \to \infty$, the conjugacy classes

$$\{\theta(E, f)\}_{f \text{ in } X(E)}$$

become equidistributed for Haar measure in the space $K^\#$ of conjugacy classes in K.

7.4 The O or SO case

(7.4.1) Let us finally consider the case in which $\mathcal{F}(w/2)$ is symplectically self–dual on C/k, and χ has order 2 (and the hypotheses of section 7.0 are in force). Then, by Poincaré duality, $\mathcal{G}((w+1)/2)$ is orthogonally self–dual as a lisse sheaf on X. The field k must have char(k) $\neq 2$, simply because χ has order 2. By hypothesis, Theorem 5.5.1 holds, so $\mathcal{G}$ has G_{geom} either SO(N) or O(N).

(7.4.2) If G_{geom} is O(N), then the lisse sheaf $\mathcal{G}((w+1)/2)$ is ι–pure of weight zero, and all its Frobenii land in G_{geom}. See Proposition 5.5.2 for various conditions which insure that G_{geom} is O(N) rather than SO(N). In particular, recall that G_{geom} is O(N) if N is odd.

(7.4.3) If G_{geom} is SO(N), we have

(7.4.3.1) $\qquad SO(N) = G_{geom} \subset G_{arith} \subset O(N)$,

where we write G_{arith} for the Zariski closure of the image of $\pi_1(X)$ under the (orthogonal) representation corresponding to $\mathcal{G}((w+1)/2)$. Thus G_{arith} is SO(N) if and only if $\det(\mathcal{G}((w+1)/2))$ is arithmetically trivial. In any case, we know that $\det(\mathcal{G}((w+1)/2))$ is of order 1 or 2, and that it is geometrically trivial (because $G_{geom} \subset SO(N)$). Thus we have

(7.4.3.2) $\qquad \det(\mathcal{G}((w+1)/2)) = \varepsilon^{\deg}$,

for $\varepsilon = \pm 1$. [So for k_2/k the quadratic extension of k inside $\bar{k}$, the pullback of $\mathcal{G}$ to $X \otimes_k k_2$ will always have $G_{arith} = G_{geom} = SO(N)$, independently of whether ε is 1 or -1.]

(7.4.4) If G_{geom} is SO(N), we can compute ε in principle as follows. If the parameter space X has a k–rational point f, then

(7.4.4.1) $\qquad \varepsilon = \det(\mathrm{Frob}_{k,f} \mid \mathcal{G}((w+1)/2))$.

If there is no k–rational point in X, there will be an E–rational point of X for any finite extension E/k of high enough degree. If we take E of **odd** degree over k, and an f in X(E), then we still have the recipe

(7.4.4.2) $\qquad \varepsilon = \det(\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2))$.

(7.4.5) If G_{geom} is SO(N) and $\varepsilon = 1$, then all Frobenii for $\mathcal{G}((w+1)/2)$ land in $G_{geom} = SO(N)$.

(7.4.6) If G_{geom} is SO(N) and ε is -1, then $G_{arith} = O(N)$ contains $G_{geom} = SO(N)$ with index two. The Frobenius conjugacy classes $\mathrm{Frob}_{E,f}$ land in $O_-(N)$ for E/k of odd degree, and they land in SO(N) for E/k of even degree.

(7.4.7) Suppose we do not know whether G_{geom} is SO or O. Here is a computational way to sort out which of the three cases

(7.4.7.1) $\qquad G_{geom} = O(N) = G_{arith}$,

(7.4.7.2) $\qquad G_{geom} = SO(N) \subset G_{arith} = O(N)$,

(7.4.7.3) $\qquad G_{geom} = SO(N) = G_{arith}$,

$\mathcal{G}((w+1)/2)$ is in. The question is whether the character of order dividing 2 of $\pi_1(X)$ given by $\det(\mathcal{G}((w+1)/2))$ is nontrivial or not, both arithmetically (i.e., on $\pi_1(X)$) and geometrically (i.e., on $\pi_1^{geom}(X)$).

Computational algorithm 7.4.8 Pick a large finite extension E/k of odd degree. For each f in X(E), compute $\det(\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2))$, which a priori is ± 1. If both 1 and -1 occur as f varies in X(E), we are in the first case 7.4.7.1. If only -1 occurs, we are in the second case 7.4.7.2. If only $+1$ occurs, we are in the third case 7.4.7.3. [The point is that in the second case we will get only -1, and in the third case we will get only $+1$, whatever the odd degree extension E/k with X(E) nonempty. If E is large, then Chebotarev for $\det(\mathcal{G}((w+1)/2))$ on X guarantees that, if we are in the first case, then both signs 1 and −1 occur as f varies over X(E).]

(7.4.9) Here is a minor variation on 7.4.8, when X(k) is nonempty.

Computational algorithm 7.4.10, when X(k) is nonempty Take a large finite extension E/k of even degree. We are in the first case 7.4.7.1 if and only if both signs occur as f varies in X(E). If only +1 occurs, then G_{geom} is SO(N), In this case, we compute ε as $\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}((w+1)/2))$ at any single k−rational point of X.

(7.4.11) Let us denote by $K \subset K_{arith}$ maximal compact subgroups of $G_{geom}(\mathbb{C})$ and of $G_{arith}(\mathbb{C})$. So we are in one of the three cases:

(7.4.11.1) $K = O(N, \mathbb{R}) = K_{arith}$,

(7.4.11.2) $K = SO(N, \mathbb{R}) \subset K_{arith} = O(N, \mathbb{R})$,

(7.4.11.3) $K = SO(N, \mathbb{R}) = K_{arith}$.

For each finite extension E/k inside $\bar{k}$, and each f in X(E), we denote by $\theta(E, f)$ the Frobenius conjugacy class in K_{arith} attached to $\mathcal{G}((w+1)/2)$ at the E−valued point f of X. Thus

(7.4.11.4) $\det(1 - T\theta(E, f)) := \iota(\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)))$

$$= \iota(\det(1 - T\alpha_k^{\deg(E/k)(w+1)}\mathrm{Frob}_{E,f} \mid H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)})))).$$

(7.4.12) If K_{arith} is $O(N, \mathbb{R})$, this equality 7.4.11.4 of characteristic polynomials determines $\theta(E, f)$ as a conjugacy class in K_{arith}. If $K_{arith} = SO(N, \mathbb{R})$, this equality of characteristic polynomials only determines $\theta(E, f)$ in $SO(N, \mathbb{R})$ up to conjugation by the ambient group $O(N, \mathbb{R})$.

(7.4.13) If $K = K_{arith}$, then by Deligne's general equidistribution theorem [De−WeII, 3.5.3], cf. also [Ka−GKM, 3.6] and [Ka−Sar, RMFEM, 9.2.6], as $\#E \to \infty$, the conjugacy classes

$$\{\theta(E, f)\}_{f \text{ in } X(E)}$$

become equidistributed for Haar measure in the space $K^\#$ of conjugacy classes in K.

(7.4.14) If $K = SO(N, \mathbb{R})$ but $K_{arith} = O(N, \mathbb{R})$, the space $O(N, \mathbb{R})^\#$ of conjugacy classes in $O(N, \mathbb{R})$ is a disjoint union

$$O_+(N, \mathbb{R})^\# \amalg O_-(N, \mathbb{R})^\#,$$

where we write $O_\varepsilon(N, \mathbb{R})^\#$ for the set of conjugacy classes of determinant ε. In this case, Deligne's general equidistribution theorem [De−WeII, 3.5.3], cf. also [Ka−Sar, RMFEM, 9.7.10], tells us that as $\#E \to \infty$ through fields E/k whose degree over k has fixed parity $\varepsilon = (-1)^{\deg(E/k)}$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X(E)}$ become equidistributed for Haar measure in the space $O_\varepsilon(N, \mathbb{R})^\#$.

(7.4.15) If $K = K_{arith} = O(N)$, the equidistribution as $\#E \to \infty$ of the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X(E)}$ in

$$O(N, \mathbb{R})^\# = O_+(N, \mathbb{R})^\# \amalg O_-(N, \mathbb{R})^\#$$

amounts to two finer statements of equidistribution. To state them, we take the Haar measure on $O(N, \mathbb{R})$ of total mass 2, restrict it to each of $O_{\pm}(N, \mathbb{R})$, and take its direct image to $O_{\pm}(N, \mathbb{R})^{\#}$. We call this "Haar measure of total mass one" on $O_{\pm}(N, \mathbb{R})^{\#}$. For each finite extension E/k, and each value of $\varepsilon = \pm 1$, denote by $X_{\varepsilon}(E)$ the subset of X(E) consisting of those points f in X(E) such that

$$\det(\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)) = \varepsilon.$$

For each choice of $\varepsilon = \pm 1$, as $\#E \to \infty$, we have

$$\#X_{\varepsilon}(E)/\#X(E) \to 1/2,$$

(by Chebotarev applied to $\det(\mathcal{G}((w+1)/2))$). Therefore, for each choice of $\varepsilon = \pm 1$, as $\#E \to \infty$ the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\varepsilon}(E)}$ become equidistributed for Haar measure of total mass one on the space $O_{\varepsilon}(N, \mathbb{R})^{\#}$.

(7.4.16) When $K = K_{arith} = O(N)$, there is another way to index the decomposition

(7.4.16.1) $$O(N, \mathbb{R})^{\#} = O_{+}(N, \mathbb{R})^{\#} \amalg O_{-}(N, \mathbb{R})^{\#}.$$

Namely, we define

(7.4.16.2) $$O_{\mathrm{sign}\ \varepsilon}(N) := \{A \text{ in } O(N) \text{ with } \det(-A) = \varepsilon\}.$$

Thus for even N there is nothing new, $O_{\mathrm{sign}\ \varepsilon}(N) = O_{\varepsilon}(N)$. But if N is odd, then $O_{\mathrm{sign}\ \varepsilon}(N) = O_{-\varepsilon}(N)$. The reason to consider this $O_{\mathrm{sign}\ \varepsilon}(N)$ decomposition is that for an orthogonal F, it is $\det(-F)$ rather than $\det(F)$ which is the sign in the functional equation.

(7.4.17) For the sake of completeness, we restate the equidistribution for this breakup (still assuming $K = K_{arith} = O(N)$). For each finite extension E/k, and each value of $\varepsilon = \pm 1$, denote by $X_{\mathrm{sign}\ \varepsilon}(E)$ the subset of X(E) consisting of those points f in X(E) such that

$$\det(-\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)) = \varepsilon.$$

For each choice of $\varepsilon = \pm 1$, as $\#E \to \infty$,

$$\#X_{\mathrm{sign}\ \varepsilon}(E)/\#X(E) \to 1/2,$$

and the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\mathrm{sign}\ \varepsilon}(E)}$ become equidistributed for the Haar measure of total mass one on the space $O_{\mathrm{sign}\ \varepsilon}(N, \mathbb{R})^{\#}$.

7.5 Interlude: a lemma on tameness and compatible systems

Lemma 7.5.1 Let k be a finite field of characteristic p, U/k a smooth, geometrically connected curve, and w an integer. Suppose for each prime $\ell \neq p$ we are given a lisse $\mathbb{Q}_{\ell}$–sheaf $\mathcal{F}_{\ell}$ on U, which is pure of weight w. Suppose the sheaves $\{\mathcal{F}_{\ell}\}_{\ell \neq p}$ form a $\mathbb{Q}$–compatible system, in the sense that for each finite extension E/k, and each point x in U(E), the characteristic polynomial

$$\det(1 - T\mathrm{Frob}_{E,x} \mid \mathcal{F}_{\ell})$$

has coefficients in $\mathbb{Q}$, independent of $\ell \neq p$. Then we have the following results.
1) All the sheaves $\mathcal{F}_\ell$ have the same rank, say r.
2) Denote by C the complete nonsingular model of U, $j : U \hookrightarrow C$ the inclusion. If for a single $\ell \neq p$ the sheaf $j_*\mathcal{F}_\ell$ is everywhere tame on C, then for every $\ell \neq p$ the sheaf $j_*\mathcal{F}_\ell$ is everywhere tame on C.
3) If $p \geq r+2$, all the sheaves $\{j_*\mathcal{F}_\ell\}_{\ell \neq p}$ are everywhere tame on C.

proof For 1), we get r as the common degree of any single characteristic polynomial of Frobenius. For 2), we use a fundamental result of Deligne [De–Const, 9.8], which tells us for each "point at infinity" y in $(C-U)(\bar{k})$, and each element γ in the inertia group I(y), the trace of the action of γ on $\mathcal{F}_\ell$ lies in $\mathbb{Z}$, independent of $\ell \neq p$. But $\mathcal{F}_\ell$ is tame at y if and only if the trace of the action of every γ in P(y) on $\mathcal{F}_\ell$ is r. For 3), we use 2) to reduce to finding a single $\ell \neq p$ for which $j_*\mathcal{F}_\ell$ is everywhere tame. Since $\mathcal{F}_\ell$ as $\mathbb{Q}_\ell$–representation of the compact group $\pi_1(U)$ admits a $\mathbb{Z}_\ell$–form, it suffices to pick an ℓ such that the profinite group $GL(r, \mathbb{Z}_\ell)$ is prime to p, or, equivalently, such that the finite group $GL(r, \mathbb{F}_\ell)$ is prime to p. The order of $GL(r, \mathbb{F}_\ell)$ is

$$\prod_{\nu=0 \text{ to } r-1} (\ell^r - \ell^\nu) = \ell^{r(r-1)/2} \times \prod_{i=1 \text{ to } r} (\ell^i - 1).$$

Take a prime ℓ whose reduction mod p is a generator of the cyclic group $\mathbb{F}_p^\times$. Since $p-1 > r$, each factor $\ell^i - 1$ is prime to p. QED

7.6 Applications to L–functions of quadratic twists of elliptic curves and of their symmetric powers over function fields

(7.6.1) We continue to work over a **finite** field k, of cardinality q and odd characteristic p. We fix a proper, smooth, geometrically connected curve C/k of genus g, and a prime number ℓ invertible in k. Over the function field k(C), we are given an elliptic curve E/k(C) with **nonconstant** j–invariant. We denote by

(7.6.1.1) $$j : U \subset C$$

the inclusion of any dense open set of C over which E/k(C) extends to an elliptic curve $\pi: \mathcal{E} \to U$.

(7.6.2) The sheaf $R^1\pi_*\bar{\mathbb{Q}}_\ell$ on U is lisse of rank 2, pure of weight one, and part of a $\mathbb{Q}$–compatible system, hence everywhere tame if $p \geq 5$. If p=3, we **assume** that $R^1\pi_*\bar{\mathbb{Q}}_\ell$ is everywhere tame. The sheaf $R^1\pi_*\bar{\mathbb{Q}}_\ell(1/2)$ on U is lisse of rank 2, pure of weight zero, and symplectically self–dual. We define

(7.6.2.1) $$\mathcal{F} := j_*R^1\pi_*\bar{\mathbb{Q}}_\ell(1/2)$$

on C. By the Néron–Ogg–Shafarevich criterion of good reduction [S–T, GR], the open set on which $\mathcal{F}$ is lisse is the largest open set over which E/k(C) has good reduction:

$$\mathrm{Sing}(\mathcal{F}) = \mathrm{Sing}(E/k(C)).$$

(7.6.3) For every integer $n \geq 0$, the lisse sheaf $\mathrm{Symm}^n(R^1\pi_*\overline{\mathbb{Q}}_\ell(1/2))$ on U is lisse of rank n+1, pure of weight zero, and everywhere tame. It is symplectically self–dual if n is odd, and it is orthogonally self–dual if n is even. Because E/k(C) has nonconstant j–invariant, the sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell(1/2)$ has $G_{\mathrm{geom}} = SL(2)$, cf. [De–WeII, 3.5.5]. This has the consequence that for every integer $n \geq 0$, $\mathrm{Symm}^n(R^1\pi_*\overline{\mathbb{Q}}_\ell(1/2))$ is geometrically irreducible (because the symmetric powers of the standard two–dimensional representation of SL(2) are irreducible). For odd (respectively even) n, $\mathrm{Symm}^n(R^1\pi_*\overline{\mathbb{Q}}_\ell(1/2))$ is symplectically (respectively orthogonally) self–dual.

(7.6.4) For every integer $n \geq 0$, we define a geometrically irreducible middle extension sheaf $\mathcal{F}_n$ on C by

(7.6.4.1) $$\mathcal{F}_n := j_*\mathrm{Symm}^n(R^1\pi_*\overline{\mathbb{Q}}_\ell(1/2)).$$

Thus $\mathcal{F}_1$ is the $\mathcal{F}$ defined above. For every $n \geq 0$, we have

(7.6.4.2) $$\mathrm{Sing}(\mathcal{F}_n) \subset \mathrm{Sing}(\mathcal{F}) = \mathrm{Sing}(E/k(C)).$$

(7.6.5) Suppose that at some point x in $C(\overline{k})$, the action of the local monodromy group I(x) on $\mathcal{F}$ is unipotent and nontrivial, or, equivalently [S–T, GR], that E has multiplicative reduction at x. At such a point, the action of I(x) is automatically tame (because by unipotence its image is pro–ℓ). If we pick a topological generator of the tame quotient $I^{\mathrm{tame}}(x)$ of I(x), then γ acts on $\mathcal{F}(x)$ by a single unipotent Jordan block of size two, Unip(2).

(7.6.6) At any point x in $C(\overline{k})$ where E has multiplicative reduction, a topological generator γ of $I(x)^{\mathrm{tame}}$ acts by $\mathrm{Symm}^n(\mathrm{Unip}(2)) = \mathrm{Unip}(n+1)$, a single unipotent Jordan block of size n+1. Thus we have

(7.6.6.1) $$\mathcal{F}_n(x)/\mathcal{F}_n(x)^{I(x)} \cong \mathrm{Unip}(n)$$

as representation of I(x). In particular, we have the dimension formula

(7.6.6.2) $$\dim(\mathcal{F}_n(x)/\mathcal{F}_n(x)^{I(x)}) = n.$$

Theorem 7.6.7 Let k be a finite field of odd characteristic, C/k a proper, smooth, geometrically connected curve of genus g, ℓ a prime number ℓ invertible in k, ι an embedding of $\overline{\mathbb{Q}}_\ell$ into $\mathbb{C}$. Let E/k(C) be an elliptic curve E/k(C) with nonconstant j–invariant, such that $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is everywhere tame. Let D_ν, $\nu \geq 1$, be a sequence of effective divisors in C, whose degrees $d_\nu \geq 2g+1$ are strictly increasing.

Denote by

$$j : U \subset C$$

the inclusion of any dense open set of C over which E/k(C) extends to an elliptic curve $\pi: \mathcal{E} \to U$, and put, for each $n \geq 0$,

$$\mathcal{F}_n := j_*\mathrm{Symm}^n(R^1\pi_*\overline{\mathbb{Q}}_\ell(1/2)).$$

For each pair of integers ($\nu \ge 1$, $n \ge 0$), denote

$$X_{\nu,n} := \mathit{Fct}(C, d_\nu, D_\nu, \mathrm{Sing}(\mathcal{F}_n)_{\mathrm{finite}}).$$

Denote by $\mathcal{G}_{\nu,n} := \mathrm{Twist}_{\chi_2,C,D_\nu}(\mathcal{F}_n)$ the lisse sheaf on $X_{n,\nu}$ constructed out of $\mathcal{F}_n$ and the quadratic character χ_2 of $k^\times$ by the recipe of 5.2.1, but carried out over k instead of $\overline{k}$, cf. 6.2.10.

Denote by $N_{\nu,n}$ the rank of $\mathcal{G}_{\nu,n}$. Thus

$$N_{\nu,n} \ge (2g - 2 + d_\nu)(n+1).$$

Then we have the following results.

1) Fix an even integer $n \ge 0$. Take ν sufficiently large that we have

$$d_\nu \ge 4g+4,$$

and

$$2g - 2 + d_\nu > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}_1)(\overline{k}), 72(n+1)).$$

The lisse sheaf $\mathcal{G}_{\nu,n}(1/2)$ on $X_{\nu,n}$ is ι–pure of weight zero and symplectically self–dual, and $G_{\mathrm{geom}} = Sp(N_{\nu,n})$. Put $K := USp(N_{\nu,n})$, a maximal compact subgroup of $G_{\mathrm{geom}}(\mathbb{C})$. For each finite extension E/k inside $\overline{k}$, and each f in $X_{\nu,n}(E)$, we denote by $\theta(E, f)$ the Frobenius conjugacy class in $USp(N_{\nu,n})$ attached to $\mathcal{G}_{\nu,n}(1/2)$ at the E–valued point f of $X_{\nu,n}$. Thus

$$\det(1 - T\theta(E, f)) := \iota(\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}_{\nu,n}(1/2))).$$

As $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\nu,n}(E)}$ become equidistributed for Haar measure in the space $USp(N_{\nu,n})^\#$ of conjugacy classes in $USp(N_{\nu,n})$.

2) Fix an odd integer $n \ge 0$. Suppose that for every ν, there is a $\overline{k}$–valued point in $C - D_\nu$ at which E has multiplicative reduction. Take ν sufficiently large that we have

$$d_\nu \ge 4g+4,$$

and

$$2g - 2 + d_\nu > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}_1)(\overline{k}), 72(n+1)).$$

The lisse sheaf $\mathcal{G}_{\nu,n}(1/2)$ on $X_{\nu,n}$ is ι–pure of weight zero and orthogonally self–dual, and $G_{\mathrm{geom}} = O(N_{\nu,n})$. Put $K := O(N_{\nu,n}, \mathbb{R})$, a maximal compact subgroup of $G_{\mathrm{geom}}(\mathbb{C})$. For each finite extension E/k inside $\overline{k}$, and each f in $X_{\nu,n}(E)$, we denote by $\theta(E, f)$ the Frobenius conjugacy class in $O(N_{\nu,n}, \mathbb{R})$ attached to $\mathcal{G}_{\nu,n}(1/2)$ at the E–valued point f of $X_{\nu,n}$. Thus

$$\det(1 - T\theta(E, f)) := \iota(\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}_{\nu,n}(1/2))).$$

As $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\nu,n}(E)}$ become equidistributed for Haar measure in the space $O(N_{\nu,n}, \mathbb{R})^\#$ of conjugacy classes in $O(N_{\nu,n}, \mathbb{R})$.

proof By assumption, $\mathcal{F}_1$ and hence all the sheaves $\mathcal{F}_n$ are everywhere tame. Theorem 5.5.1 will apply to $\mathcal{G}_{\nu,n}$ provided only that $d_\nu \geq 4g+4$ and

$$2g - 2 + d_\nu > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}_n)(\bar{k}), 72\mathrm{rank}(\mathcal{F}_n)).$$

Now $\mathrm{rank}(\mathcal{F}_n) = n+1$, and $\mathrm{Sing}(\mathcal{F}_n) \subset \mathrm{Sing}(\mathcal{F}_1)$, so this last inequality will hold if

$$2g - 2 + d_\nu > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}_1)(\bar{k}), 72(n+1)).$$

Assertion 1) is thus an instance of the Sp case 7.3 of the preceding discussion. In assertion 2), the hypothesis of multiplicative reduction at a point x of $C - D_\nu$ gives

$$\dim(\mathcal{F}_n(x)/\mathcal{F}_n(x)^{I(x)}) = n.$$

As n is odd, Proposition 5.5.2, part 1) shows that G_{geom} is $O(N_{\nu,n})$ rather than $SO(N_{\nu,n})$. Once we have this, assertion 2) becomes an instance of the $G_{geom} = O = G_{arith}$ case 7.4.15 of the preceding discussion. QED

7.7 Applications to L–functions of Prym varieties

Theorem 7.7.1 Let k be a finite field of odd characteristic, C/k a proper, smooth, geometrically connected curve of genus g, ℓ a prime number ℓ invertible in k, ι an embedding of $\bar{\mathbb{Q}}_\ell$ into $\mathbb{C}$. Let D be an effective divisor in C, whose degree d satisfies

$$d \geq 4g+4$$

and

$$2g - 2 + d > 4.$$

Take $\mathcal{F}$ to be the constant sheaf $\bar{\mathbb{Q}}_\ell$ on C. Thus $\mathcal{F}$ is everywhere lisse of rank one, pure of weight zero, and orthogonally self–dual.

Denote

$$X := \mathit{Fct}(C, d, D, \emptyset).$$

Denote by $\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\bar{\mathbb{Q}}_\ell)$ the lisse sheaf on X constructed out of $\mathcal{F} := \bar{\mathbb{Q}}_\ell$ and the quadratic character χ_2 of $k^\times$ by the recipe of 5.2.1, but carried out over k instead of $\bar{k}$, cf. 6.2.10. Concretely, for E/k a finite extension of k, and f in X(E), the stalk $\mathcal{G}_f$ of $\mathcal{G}$ at f is $H^1(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi_2(f)})$, the H^1 of the Prym variety attached to the double cover $C(f^{1/2})$ of $C\otimes_k E$, or, equivalently, to the odd part of $H^1(C(f^{1/2})\otimes_E\bar{k}, \bar{\mathbb{Q}}_\ell)$.

Denote by N the rank of $\mathcal{G}$. Thus

$$N \geq 2g - 2 + d.$$

Then the lisse sheaf $\mathcal{G}(1/2)$ on X is ι–pure of weight zero and symplectically self–dual, and $G_{geom} = Sp(N)$. Put $K := USp(N)$, a maximal compact subgroup of $G_{geom}(\mathbb{C})$. For each finite extension E/k inside $\bar{k}$, and each f in X(E), denote by $\theta(E, f)$ the Frobenius conjugacy class in USp(N) attached to $\mathcal{G}(1/2)$ at the E–valued point f of X. Thus

$$\det(1 - T\theta(E, f)) := \iota(\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}(1/2)))$$
$$= \iota(\det(1 - T\mathrm{Frob}_E \mid H^1{}_c(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi_2(f)})(1/2))).$$

As $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X(E)}$ become equidistributed for Haar measure in the space $USp(N)^\#$ of conjugacy classes in USp(N).

proof This is a special case of the Sp discussion 7.3 above. QED

7.8 Families of hyperelliptic curves as a special case

(7.8.1) If the curve C is $\mathbb{P}^1$, then the Prym variety attached to the double cover $C(f^{1/2})$ of $C\otimes_k E$ is simply the Jacobian of the hyperelliptic curve of equation $y^2 = f(x)$. So the sheaf $\mathcal{G}$ in this case is just the H^1 along the fibres in the family of hyperelliptic curves $\{y^2 = f(x)\}_{f \text{ in } X}$ over the space $X := Fct(\mathbb{P}^1, d, D, \emptyset)$. As $\mathbb{P}^1$ has genus g=0, we find that $\mathcal{G}$ has G_{geom} the full symplectic group, provided only that the effective D has degree $d \geq 7$. If we successively take for D the divisor $d\infty$, d = 7, 8, 9, ..., we recover [Ka–Sar, RMFEM, 10.1.18.3 and 10.1.18.5] in every genus $g \geq 3$.

7.9 Application to L–functions of χ–components of Jacobians of cyclic coverings of degree $n \geq 3$ in odd characteristic

Theorem 7.9.1 Let k be a finite field of odd characteristic p, C/k a proper, smooth, geometrically connected curve of genus g, ℓ a prime number ℓ invertible in k, ι an embedding of $\overline{\mathbb{Q}}_\ell$ into $\mathbb{C}$. Let

$$\chi : k^\times \to \overline{\mathbb{Q}}_\ell{}^\times$$

be a nontrivial character of $k^\times$, of order $n \geq 3$. Define

$$m := \text{the order of } \chi\times\chi_2.$$

[Thus if n is odd, m = 2n, if n is 2d with d odd then m = d, and if n is divisble by 4 then m = n.]

Let D be an effective divisor in C, whose degree d satisfies

$$d \geq 4g+4,$$

and

$$2g - 2 + d > 4.$$

Take $\mathcal{F}$ to be the constant sheaf $\overline{\mathbb{Q}}_\ell$ on C. Thus $\mathcal{F}$ is everywhere lisse of rank one, and pure of weight zero.

Denote

$$X := Fct(C, d, D, \emptyset).$$

Denote by $\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\overline{\mathbb{Q}}_\ell)$ the lisse sheaf on X constructed out of $\mathcal{F} := \overline{\mathbb{Q}}_\ell$ and the character

χ of $k^\times$ by the recipe of 5.2.1, but carried out over k instead of $\bar{k}$, cf. 6.2.10. Concretely, for E/k a finite extension of k, and f in X(E), the stalk $\mathcal{G}_f$ of $\mathcal{G}$ at f is

$$\mathcal{G}_f = H^1_c(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi(f)}),$$

the χ–component of $H^1(C(f^{1/n})\otimes_E\bar{k}, \bar{\mathbb{Q}}_\ell)$.

Denote by N the rank of $\mathcal{G}$. Thus

$$N \geq 2g - 2 + d.$$

Suppose further that one of the following three conditions is satisfied:

a) n is odd,

b) $n \equiv 0 \bmod 4$,

c) n is even, n/2 is odd, and over $\bar{k}$, $D = \sum a_iP_i$ with each a_i odd.

Then the lisse sheaf $\mathcal{G}$ on X is ι–pure of weight one, and G_{geom} is the group

$$GL_m(N) := \{A \text{ in } GL(N) \mid \det(A)^m = 1\}.$$

proof Write D as the sum of effective divisors $D_1 + D_2$ with degrees $d_1 \geq 2g+2$ and $d_2 \geq 2g+1$, such that $D_2 = \sum c_iP_i$ has all its nonzero c_i invertible in k. This is possible by Corollary 5.4.8, part 2). If g=0, do this so that $d_2 \geq 2$. (If g = 0, then $d \geq 6 > 4g+5$, so we may apply Corollary 5.4.8, part 1).)

Pick f_1 and f_2 as in the statement of Theorem 5.4.9. Then the pullback

$$\mathcal{H} := [t \mapsto f_1(t-f_2)]^*\mathcal{G}$$

of $\mathcal{G}_\nu$ to $\mathbb{A}^1$ – CritVal(f_2, $\mathcal{F}\otimes\mathcal{L}_{\chi(f_1)}$) has G_{geom} containing SL(N). Moreover, f_2 has at least one critical value, and the local monodromy of $\mathcal{H}$ at each critical value of f_2 is a pseudoreflection of determinant $\chi\times\chi_2$, a character of order m. The local monodromy of $\mathcal{H}$ at the image under f_2 of each zero of f_1 is a pseudoreflection of determinant χ^2, a character of order dividing m. The sheaf $\mathcal{H}$ has no other finite singularities, and is tame at ∞. Therefore det($\mathcal{H}$) as a character of $\pi_1{}^{geom}$ is generated by its local monodromies at finite distance, so has order m. Since $n \geq 3$, we have $m \geq 3$. By the paucity of choice, G_{geom} for $\mathcal{H}$ is $GL_m(N)$.

Therefore G_{geom} for $\mathcal{G}$ itself contains $GL_m(N)$. So it suffices to show that we have an a priori inclusion $G_{geom} \subset GL_m(N)$, i.e., to prove the following lemma.

Lemma 7.9.2 Hypotheses and notations as in Theorem 7.9.1 above, $\det(\mathcal{G})^{\otimes m}$ is geometrically trivial.

proof Suppose first that either a) or b) holds. Then m is the number of roots of unity in the field $\mathbb{Q}(\chi)$, and the result follows from the fact that $\mathcal{G}$ is part of a $\mathbb{Q}(\chi)$–compatible system of lisse sheaves on X, cf. [Ka–ACT, the "trivial" part of the proof of 5.2 bis].

Suppose now that c) holds. Then $n = 2m$ with m odd. All the a_i are nonzero mod n, because they are all odd. The idea is to use the argument of [Ka–ACT, 5.2 bis]. The sheaf $\mathcal{G}$ was constructed as the image of the natural "forget supports" map

$$\mathcal{G}_!(\chi) := R^1\pi_!(\mathcal{L}_{\chi(f)}) \to R^1\pi_*(\mathcal{L}_{\chi(f)}) := \mathcal{G}_*(\chi).$$

Because all the a_i are nonzero mod n, this map is an isomorphism, as one verifies by checking fibre by fibre. In other words, we have

$$\mathcal{G}_!(\chi) \cong \mathcal{G}(\chi).$$

So it suffices to show that $\det(\mathcal{G}_!(\chi))^{\otimes m}$ is geometrically constant.

If we replace χ by the quadratic character χ_2 of $k^\times$, and form the analogous sheaves $\mathcal{G}_!(\chi_2)$ and $\mathcal{G}(\chi_2)$, we have

$$\mathcal{G}_!(\chi_2) \cong \mathcal{G}(\chi_2),$$

because all the a_i are odd. But $\mathcal{G}(\chi_2)$ is symplectic, so $\det(\mathcal{G}(\chi_2))$ and hence $\det(\mathcal{G}_!(\chi_2))$ are geometrically trivial. So it suffices to show that we have a geometric isomorphism

$$\det(\mathcal{G}_!(\chi))^{\otimes m} \cong \det(\mathcal{G}_!(\chi_2))^{\otimes m}.$$

This results from the "change of λ, reduction mod λ, change of χ" argument of [Ka–ACT, 5.2 bis], which is valid independent of any hypotheses on the a_i. QED

(7.9.3) What happens in Theorem 7.9.1 above if we allow χ to have order $n \geq 3$ with $n = 2m$ with m odd, $m \geq 3$, but do not make any hypothesis on D? The order of $\chi\times\chi_2$ is m, but $\mathbb{Q}(\chi)$ contains $n = 2m$ roots of unity. The compatible system argument of [Ka–ACT, the "trivial" part of the proof of 5.2 bis] shows that $\det(\mathcal{G})^{\otimes 2m}$ is geometrically trivial. The argument in the proof of Theorem 7.9.1 concerning $\mathcal{H}$ remains valid, and shows that $\det(\mathcal{H})$ has geometric order m. Thus G_{geom} for $\mathcal{G}$ is either $GL_m(N)$ or it is $GL_{2m}(N)$. In fact, both cases arise. Here is the precise result.

Theorem 7.9.4 Notations as in Theorem 7.9.1, suppose that χ has order $n \geq 3$ with $n = 2m$ and m odd, $m \geq 3$. If there exists an index i such that a_i is even but not divisible by n, then G_{geom} for

$$\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\overline{\mathbb{Q}}_\ell)$$

is $GL_{2m}(N)$. If there exists no such index i, i.e., if every a_i is either odd or divisible by n, then G_{geom} for $\mathcal{G}$ is $GL_m(N)$.

proof Since χ is of order $n \geq 3$, we know already that G_{geom} for $\mathcal{G}$ is either $GL_m(N)$ or it is $GL_{2m}(N)$. We need only determine whether or not $\det(\mathcal{G})^{\otimes m}$ is geometrically trivial.

Because we are trying to determine G_{geom}, we may extend scalars from k to any finite extension E/k (and simultaneously replace χ by the character $\chi\circ\mathrm{Norm}_{E/k}$). Thus it suffices to treat universally the case in which $D = \sum a_i P_i$ with each P_i a k−valued point of C. Moreover, we know that $\det(\mathcal{G})^{\otimes m}$ has, geometrically, order either one or two. We may and will further assume that k is large enough that, in addition, both of the following conditions hold:

1) $\det(\mathcal{G})^{\otimes m}$ is geometrically trivial if and only if $\det(\mathcal{G})^{\otimes m}$ is constant on the set of k−valued points f in X(k).

2) $\#X(k)/\#L(D) > 1/2$.

For a nontrivial character ρ of $k^\times$, of order denoted order(ρ), denote by $\mathrm{Div}(\rho) \subset D^{\mathrm{red}}$ the set of those points P_i whose coefficient a_i is divisible by order(ρ).

Given f in X(k), we have the sheaf $\mathcal{L}_{\rho(f)}$ on (C−D)[1/f]. Denote by

$$j : (C-D)[1/f] \to C$$

the inclusion. We have a short exact sequence of sheaves on C,

$$0 \to j_!\mathcal{L}_{\rho(f)} \to j_*\mathcal{L}_{\rho(f)} \to \oplus_{P_i \text{ in } \mathrm{Div}(\rho)} (j_*\mathcal{L}_{\rho(f)})|P_i \to 0.$$

At each P_i in Div(ρ), $(j_*\mathcal{L}_{\rho(f)})|P_i$ is a skyscraper sheaf of rank one at P_i, on which Frob_{k,P_i} acts as a scalar. This scalar is computed in terms of the auxiliary choice of a uniformizing parameter π_i at P_i as follows. In the local ring $\mathcal{O}_{C,P_i}$, one forms the unit

$$f_i := f\times(\pi_i)^{a_i}.$$

Its value $f_i(P_i)$ in the residue field k is nonzero (because f has a pole of order a_i at P_i) and it is well defined in $k^\times/(k^\times)^{\mathrm{order}(\rho)}$, independent of the auxiliary choice of π_i (because $a_i \equiv 0 \bmod \mathrm{order}(\rho)$). Then we have

$$\mathrm{Frob}_{k,P_i} \mid (j^*\mathcal{L}_{\rho(f)})|P = \rho(f_i(P_i)).$$

The long exact cohomology sequence gives a short exact sequence

$$0 \to H^0(C\otimes_k\bar{k}, \oplus_{P_i \text{ in } \mathrm{Div}(\rho)} (j_*\mathcal{L}_{\rho(f)})|P_i)$$
$$\to H^1(C\otimes_k\bar{k}, j_!\mathcal{L}_{\rho(f)}) \to H^1(C\otimes_k\bar{k}, j_*\mathcal{L}_{\rho(f)}) \to 0.$$

This in turn gives

$$\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}_!(\rho)) = \det(\mathrm{Frob}_{k,f} \mid \mathcal{G}(\rho))(\textstyle\prod_{P_i \text{ in } \mathrm{Div}(\rho)} \rho(f_i(P_i))).$$

Now take ρ to be successively χ and the quadratic character, χ_2. We obtain

$$\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}_!(\chi)) = \det(\mathrm{Frob}_{k,f} \mid \mathcal{G}(\chi))(\textstyle\prod_{P_i \text{ in } \mathrm{Div}(\chi)} \chi(f_i(P_i)))$$

and

$$\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}_!(\chi_2)) = \det(\mathrm{Frob}_{k,f} \mid \mathcal{G}(\chi_2))(\textstyle\prod_{P_i \text{ in } \mathrm{Div}(\chi_2)} \chi_2(f_i(P_i))).$$

Raise each of these relations to the m'th power, and remember that $\chi^m = (\chi_2)^m = \chi_2$. We get

$$\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}_!(\chi_2))^m/\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}_!(\chi))^m$$
$$= \det(\mathrm{Frob}_{k,f} \mid \mathcal{G}(\chi_2))^m/\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}(\chi))^m \times (\prod_{P_i \text{ in } \mathrm{Div}(\chi_2) - \mathrm{Div}(\chi)} \chi_2(f_i(P_i))).$$

We have already remarked above that $\det(\mathcal{G}_!(\chi_2))^{\otimes m}/\det(\mathcal{G}_!(\chi))^{\otimes m}$ is geometrically constant, so the left hand side is a constant function of f in X(k). As $\mathcal{G}(\chi_2)$ is symplectic, the factor

$$\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}(\chi_2))^m$$

is a constant function of f in X(k). Thus

$$\det(\mathrm{Frob}_{k,f} \mid \mathcal{G}(\chi))^m/(\prod_{P_i \text{ in } \mathrm{Div}(\chi_2) - \mathrm{Div}(\chi)} \chi_2(f_i(P_i)))$$

is a constant function of f in X(k). Therefore $\det(\mathcal{G}(\chi))^{\otimes m}$ is geometrically constant if and only if the expression

$$(\prod_{P_i \text{ in } \mathrm{Div}(\chi_2) - \mathrm{Div}(\chi)} \chi_2(f_i(P_i))) = \chi_2(\prod_{P_i \text{ in } \mathrm{Div}(\chi_2) - \mathrm{Div}(\chi)} f_i(P_i))$$

is a constant function of f in X(k).

The set $\mathrm{Div}(\chi_2) - \mathrm{Div}(\chi)$ consists precisely of those P_i in D such that a_i is even but not divisible by n. If this set is empty, then $\det(\mathcal{G}(\chi))^m$ is geometrically constant, as required.

Suppose that

$$E := \mathrm{Div}(\chi_2) - \mathrm{Div}(\chi)$$

is nonempty. We must show that as f varies in X(k), the expression

$$(\prod_{P_i \text{ in } E} \chi_2(f_i(P_i)))$$

is not constant. Equivalently, we must show that as f varies in X(k), the product $\prod_{P_i \text{ in } E} f_i(P_i)$ in $k^\times$ assumes both square and nonsquare values. If #E is odd, this is easy to see. Indeed, given f in X(k), consider also αf, for f in $k^\times$ which is a nonsquare. If #E is even, we must work a bit harder. Here is an argument which works irrespective of the cardinality of E, but just requires E to be nonempty.

Each P_i in E has multiplicity a_i in D which is **even** (and nonzero mod n). In particular, for each P_i in E, we have

$$a_i - 1 \geq a_i/2.$$

Thus we have

$$\deg(D - E) \geq \deg(D)/2 \geq 2g+2 > 2g-2.$$

Now consider the map

$$L(D) \to \prod_{P_i \text{ in } E} k,$$
$$f \text{ in } L(D) \to \prod_{P_i \text{ in } E} f_i(P_i).$$

This is a linear map, whose kernel is L(D − E). So we have a left exact sequence

$$0 \to L(D - E) \to L(D) \to \prod_{P_i \text{ in } E} k.$$

Since both D and D−E have degree > 2g−2, a dimension count shows that the last map is surjective:

$$L(D) \twoheadrightarrow \prod_{P_i \text{ in } E} k.$$

Let us denote by $L(D)(\times)$ the subset of $L(D)$ which, under the above map, lands in $\prod_{P_i \text{ in } E} k^\times$. Thus

$$L(D)(\times) = L(D) - \cup_{P \text{ in } E} L(D - P),$$
$$L(D)(\times) \twoheadrightarrow \prod_{P_i \text{ in } E} k^\times.$$

We next restrict this last map to $X(k) \subset L(D)(\times)$.

$$X(k) \to \prod_{P_i \text{ in } E} k^\times.$$

Suppose that for every f in X(k), $\prod_{P_i \text{ in } E} f_i(P_i)$ is a square [resp. a nonsquare] in k. Denote by Γ the subset of $\prod_{P_i \text{ in } E} k^\times$ consisting of those tuples whose product is a square [resp. a nonsquare]. For each γ in Γ, denote by $X(k)(\gamma)$ its inverse image in X(k). Then $X(k)(\gamma)$ lies in $L(D)(\gamma)$, the inverse image of γ in L(D). Now $L(D)(\gamma)$ is an additive torsor under L(D−E), so it has cardinality that of L(D−E). So we have a trivial inequality

$$\#X(k)(\gamma) \le \#L(D)(\gamma) = \#L(D-E).$$

Summing over Γ, which has cardinality $(1/2)(q-1)^{\#E}$, we find

$$\begin{aligned}\#X(k) = \textstyle\sum_\gamma \#X(k)(\gamma) &\le \#L(D-E)\#\Gamma\\ &\le \#L(D)\times q^{-\#E}\times(1/2)(q-1)^{\#E}\\ &\le (1/2)\#L(D)((q-1)/q)^{\#E} < (1/2)\#L(D).\end{aligned}$$

This inequality contradicts the assumption that k was large enough that $\#X(k)/\#L(D) > 1/2$. Therefore the expression

$$\left(\prod_{P_i \text{ in } \mathrm{Div}(\chi_2) - \mathrm{Div}(\chi)} \chi_2(f_i(P_i))\right)$$

assumes both values ± 1 as f varies over X(k). This in turn shows that $\det(\mathcal{G}(\chi))^{\otimes m}$ is not geometrically constant. QED

(7.9.4) We now wish to give explicit equidistribution results for sheaves

$$\mathcal{G}(\chi) := \mathrm{Twist}_{\chi,C,D}(\overline{\mathbb{Q}}_\ell)$$

on X constructed above, χ of order $n \ge 3$. We have determined that G_{geom} for $\mathcal{G}(\chi)$ is of the form $GL_\nu(N)$, with ν usually equal to m := the order of $\chi\times\chi_2$, but sometimes ν can be n. We know that $\nu = m$ except in the case that $n = 2m$ with m odd, $m \ge 3$, and, over $\overline{k}$, D is $\sum a_i P_i$ with each a_i either odd or divisible by n, in which case $\nu = n$.

Arithmetic Determinant Formula 7.9.5 Let k be a finite field of odd characteristic p, C/k a proper, smooth, geometrically connected curve of genus g, ℓ a prime number ℓ invertible in k, ι an embedding of $\overline{\mathbb{Q}}_\ell$ into $\mathbb{C}$. Let

$$\chi : k^\times \to \overline{\mathbb{Q}}_\ell{}^\times$$

be a nontrivial character of $k^\times$, of order n.

Let D be an effective divisor in C, whose degree d satisfies

$$d \geq 4g+4,$$

and

$$2g - 2 + d > 4.$$

Suppose that, over $\overline{k}$, D is $\sum a_i P_i$. Consider the following product of Gauss sums:

$$\mathrm{Const}(\chi, D) := \varepsilon(\chi_2, D) q^{g-1}(-G(\psi, \chi))^d (\prod_i (-G(\psi, \chi^{-a_i}))).$$

Here ψ is any nontrivial $\overline{\mathbb{Q}}_\ell$–valued additive character of k, and we define, for any $\overline{\mathbb{Q}}_\ell$–valued character ρ of $k^\times$, possibly trivial,

$$G(\psi, \rho) := \sum_{x \text{ in } k^\times} \psi(x)\rho(x).$$

[Thus $G(\psi, \mathbb{1}) = -1$.] The quantity $\varepsilon(\chi_2, D)$ is defined to be

$$\varepsilon(\chi_2, D) = \chi_2(-1)^S, \text{ for}$$

$$S := (1/2)(\textstyle\sum_{i \text{ with } a_i \text{ even}} a_i) + (1/2)(\sum_{i \text{ with } a_i \text{ odd}} (1 + a_i)).$$

Equivalently, $\varepsilon(\chi_2, D)$ is that choice of ± 1 such that

$$\mathrm{Const}(\chi_2, D) = \text{an integer power of q}.$$

The quantity $\mathrm{Const}(\chi, D)$ lies in $\mathbb{Q}(\chi)$, and does not depend on the auxiliary choice of ψ used to define it (because $d = \sum a_i$).

Take $\mathcal{F}$ to be the constant sheaf $\overline{\mathbb{Q}}_\ell$ on C. Thus $\mathcal{F}$ is everywhere lisse of rank one, and pure of weight zero.

Denote

$$X := \mathit{Fct}(C, d, D, \emptyset).$$

Denote by

$$\mathcal{G}(\chi) := \mathrm{Twist}_{\chi,C,D}(\overline{\mathbb{Q}}_\ell)$$

the lisse sheaf on X constructed out of $\mathcal{F} := \overline{\mathbb{Q}}_\ell$ and the character χ of $k^\times$ by the recipe of 5.2.1, but carried out over k instead of $\overline{k}$, cf. 6.2.10. Concretely, for E/k a finite extension of k, and f in X(E), the stalk $\mathcal{G}_f$ of $\mathcal{G}$ at f is $H^1(C \otimes_k \overline{k}, j_* \mathcal{L}_{\chi(f)})$, the χ–component of $H^1(C(f^{1/n}) \otimes_E \overline{k}, \overline{\mathbb{Q}}_\ell)$.

Denote by

$$\nu := \text{the geometric order of } \det(\mathcal{G}(\chi)).$$

Then we have the following arithmetic determinant formula.

$$\det(\mathcal{G}(\chi))^{\otimes \nu} = \beta^{\deg} \text{ for } \beta = \mathrm{Const}(\chi, D)^\nu.$$

proof If χ is χ_2, then $\mathcal{G}(\chi_2)(1/2)$ is symplectic, and pure of weight zero, so $\nu = 1$, the rank of

$\mathcal{G}(\chi_2)$ is even, and $\det(\mathcal{G}(\chi_2))$ is given by $\beta^{\deg}$ for $\beta = q^{\mathrm{rank}(\mathcal{G}(\chi_2))/2}$. So the assertion is correct in this case.

If χ has order $n \geq 3$, then what we are asserting is that for every finite extension E/k, and every f in X(E), the ratio

$$\det(\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi(f)}))/\mathrm{Const}(\chi, D)^{\deg(E/k)}$$

is a root of unity of order dividing ν.

Let us first treat the easy case, in which ν is the number of roots of unity in the field $\mathbb{Q}(\chi)$. Since both numerator and denominator lie in $\mathbb{Q}(\chi)$, we need only show that the ratio is a root of unity. So we may replace the numerator by

$$\det(-\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi(f)})),$$

which is the reciprocal of the constant in the functional equation for the L–function of $C\otimes_k E$ with coefficients in $j_*\mathcal{L}_{\chi(f)}$. This is an abelian L–function, with everywhere tame character, and its constant is a product of usual Gauss sums, as explained in Tate's thesis, cf. [De–Const, 5.9 and 5.10]. By using the Hasse–Davenport theorem to control the behavior of $-G(\psi, \rho)$ under field extension, it is an elementary exercise to check that, up to roots of unity, the reciprocal of our $\mathrm{Const}(\chi, D)^{\deg(E/k)}$ agrees with the global constant for the L–function of $C\otimes_k E$ with coefficients in $j_*\mathcal{L}_{\chi(f)}$.

The harder case is that in which $n = 2m$ with m odd, and every a_i either odd or divisible by n. Here we must show that for every finite extension E/k, and every f in X(E), the ratio

$$\det(\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi(f)}))^m/\mathrm{Const}(\chi, D)^{m\deg(E/k)},$$

a priori ± 1 by the above argument, is in fact 1 rather than -1. Let us denote by D_{odd} the set of P_i in D whose a_i is odd. Then the points in D but not in D_{odd} all have their a_i divisible by n. Let us denote

$$U := (C - D_{\mathrm{odd}})\otimes_k E - (\text{zeroes of } f).$$

So for ρ any character of $k^\times$ of the form

$$\chi_2 \times (\text{a character of order dividing } m),$$

$j_*\mathcal{L}_{\rho(f)}$ is a lisse rank one sheaf on U, extended by zero to all of $C\otimes_k E$. For any finite extension E_1/E, and any E_1–valued point P in $D^{\mathrm{red}} - D_{\mathrm{odd}}$, i.e., a point of $C\otimes_k E_1$ at which f has a pole of order divisible by n, pick a uniformizing parameter π at P, and define $\bar{f}(P)$ in $E_1^\times$ to be the reduction mod π of the π–adic unit $f/\pi^{\mathrm{ord}_P(f)}$. Then $\bar{f}(P)$ is well defined in $(E_1)^\times/(\text{n'th powers})$ independent of the auxiliary choice of uniformizing parameter π, and $\mathrm{Frob}_{E_1,P} \mid j_*\mathcal{L}_{\rho(f)}$ is the scalar

$$\mathrm{Frob}_{E_1,P} \mid j_*\mathcal{L}_{\rho(f)} = \rho(\mathrm{Norm}_{E_1/k}(\bar{f}(P))).$$

Thus, for any such ρ we have

$$H^*(C\otimes_k\bar{k}, j_*\mathcal{L}_{\rho(f)}) = H^*_c(U\otimes_E\bar{k}, j_*\mathcal{L}_{\rho(f)}),$$

and these groups vanish for $i \neq 1$.

The idea is to show that for all such ρ, the ratio

$$\mathrm{Ratio}(\rho) := \det(\mathrm{Frob}_E \mid H^1_c(U\otimes_E\bar{k}, j_*\mathcal{L}_{\rho(f)}))^m/\mathrm{Const}(\rho, D)^{m\deg(E/k)},$$

a priori ±1, is in fact 1. We proceed by induction on the number of distinct odd primes dividing the order of χ. If there are none, then χ is χ_2 and we are done. In carrying out the induction, we have $\chi = \rho\times\Lambda$, with Λ a character of some odd ℓ–power order, and ρ of order prime to ℓ. We then pick a finite place $\lambda|\ell$ of $\mathbb{Q}(\chi)$. As $\mathbb{Z}[\chi]$–valued functions on $k^\times$, $\rho \equiv \rho\times\Lambda \bmod \lambda$. In Ratio($\rho$) and in Ratio($\rho\times\Lambda$), both numerator and denominator are λ–adic units, and we have congruences

$$\det(\mathrm{Frob}_E \mid H^1_c(U\otimes_E\bar{k}, j_*\mathcal{L}_{\rho(f)}))$$
$$\equiv \det(\mathrm{Frob}_E \mid H^1_c(U\otimes_E\bar{k}, j_*\mathcal{L}_{(\rho\times\Lambda)(f)})) \bmod \lambda,$$

and

$$\mathrm{Const}(\rho, D) \equiv \mathrm{Constr}(\rho\times\Lambda, D) \bmod \lambda.$$

So we find a congruence

$$\mathrm{Ratio}(\rho) \equiv \mathrm{Ratio}(\rho\times\Lambda) \bmod \lambda.$$

Since both ratios are ±1, we infer that we have an equality

$$\mathrm{Ratio}(\rho) = \mathrm{Ratio}(\rho\times\Lambda).$$

Proceeding in this way, we eventually get Ratio(χ) = Ratio(χ_2). QED

Explicit Equidistribution Corollary 7.9.6 Hypotheses and notations as in 7.9.5 above, suppose χ has order $n \geq 3$. Denote by N the rank of $\mathcal{G}(\chi)$, and by ν the geometric order of $\det(\mathcal{G}(\chi))$. Put

$$U_\nu(N) := \{A \text{ in } U(N) \mid \det(A)^\nu = 1\},$$

a maximal compact subgroup of $G_{geom}(\mathbb{C})$. Denote by α any N'th root of $1/\mathrm{Const}(\chi, D)$. Then $\mathcal{G}\otimes(\alpha)^{\deg}$ is pure of weight zero, and has $G_{arith} = G_{geom}$. For each finite extension E/k inside $\bar{k}$, and each f in X(E), we denote by $\theta(E, f)$ the Frobenius conjugacy class in $U_\nu(N)$ attached to $\mathcal{G}\otimes(\alpha)^{\deg}$ at the E–valued point f of X_ν. Thus

$$\det(1 - T\theta(E, f)) := \iota(\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}\otimes(\alpha)^{\deg}))$$
$$= \iota(\det(1 - (\alpha)^{\deg(E/k)}T\mathrm{Frob}_E \mid H^1_c(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi(f)}))).$$

As $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X(E)}$ become equidistributed for Haar measure in the space $U_\nu(N)^\#$ of conjugacy classes in $U_\nu(N)$.

7.10 Application to L–functions of χ–components of Jacobians of cyclic coverings of odd degree $n \geq 3$ in characteristic 2

(7.10.1) The results in this case are very similar to those we found above in odd characteristic.

Theorem 7.10.2 Let k be a finite field of characteristic 2, C/k a proper, smooth, geometrically connected curve of genus g, ℓ a prime number ℓ invertible in k, ι an embedding of $\overline{\mathbb{Q}}_\ell$ into $\mathbb{C}$. Let

$$\chi : k^\times \to \overline{\mathbb{Q}}_\ell{}^\times$$

be a nontrivial character of $k^\times$, of (necessarily odd) order $n \geq 3$.

Let D be an effective divisor in C, whose degree d satisfies

$$d \geq 12g + 7$$

(and hence $2g - 2 + d > 4$ automatically). Over $\overline{k}$, write D as $\sum_i a_i P_i$.

Take $\mathcal{F}$ to be the constant sheaf $\overline{\mathbb{Q}}_\ell$ on C. Thus $\mathcal{F}$ is everywhere lisse of rank one, and pure of weight zero.

Denote

$$X := \mathit{Fct}(C, d, D, \emptyset).$$

Denote by

$$\mathcal{G} := \mathrm{Twist}_{\chi,C,D}(\overline{\mathbb{Q}}_\ell)$$

the lisse sheaf on X constructed out of $\mathcal{F} := \overline{\mathbb{Q}}_\ell$ and the character χ of $k^\times$ by the recipe of 5.2.1, but carried out over k instead of $\overline{k}$, cf. 6.2.10. Concretely, for E/k a finite extension of k, and f in $X_\gamma(E)$, the stalk $\mathcal{G}_f$ of $\mathcal{G}$ at f is $H^1_c(C\otimes_k\overline{k}, j_*\mathcal{L}_{\chi(f)})$, the χ–component of $H^1(C(f^{1/n})\otimes_E\overline{k}, \overline{\mathbb{Q}}_\ell)$.

Denote by N the rank of $\mathcal{G}$. Thus

$$N \geq 2g - 2 + d.$$

Then the lisse sheaf $\mathcal{G}$ on X is ι–pure of weight one, and G_{geom} is the group

$$GL_{2n}(N) := \{A \text{ in } GL(N) \mid \det(A)^{2n} = 1\}.$$

Define

$$\mathrm{Const}(\chi, D) := q^{g-1}(-G(\psi, \chi))^d \prod_i (-G(\psi, \chi^{-a_i})).$$

Denote by α any N'th root of $1/\mathrm{Const}(\chi, D)$. Then $\mathcal{G}\otimes(\alpha)^{\deg}$ is pure of weight zero, and has $G_{arith} = G_{geom}$. For each finite extension E/k inside $\overline{k}$, and each f in X(E), we denote by $\theta(E, f)$ the Frobenius conjugacy class in $U_{2n}(N)$ attached to $\mathcal{G}\otimes(\alpha)^{\deg}$ at the E–valued point f of X_γ. Thus

$$\det(1 - T\theta(E, f)) := \iota(\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}\otimes(\alpha)^{\deg}))$$
$$= \iota(\det(1 - (\alpha)^{\deg(E/k)}T\mathrm{Frob}_E \mid H^1_c(C\otimes_k\overline{k}, j_*\mathcal{L}_{\chi(f)}))).$$

As $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X(E)}$ become equidistributed for Haar measure in the space $U_{2n}(N)^\#$ of conjugacy classes in $U_{2n}(N)$.

proof That G_{geom} for $\mathcal{G}$ contains SL(N) is a special case of Theorem 5.7.1. Because $\mathcal{G}$ is part of a $\mathbb{Q}(\chi)$–compatible system, and 2n is the number of roots of unity in $\mathbb{Q}(\chi)$, $\det(\mathcal{G})^{\otimes 2n}$ is geometrically trivial, and hence G_{geom} lies in $GL_{2n}(N)$. To show that G_{geom} contains $GL_{2n}(N)$, we argue as follows. Exactly as in the proof of Theorem 5.6.2, a pullback $\mathcal{H}$ of $\mathcal{G}$ to $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$ has local monodromy at each critical value of f_2 a pseudoreflection whose determinant has order 2n, and G_{geom} for $\mathcal{H}$ contains SL(N). Therefore G_{geom} for $\mathcal{H}$ contains $GL_{2n}(N)$, and hence G_{geom} for $\mathcal{G}$ contains $GL_{2n}(N)$.

Exactly as in the proof of 7.9.5, Tate's theory of local constants for abelian L–functions shows that we have an isomorphism $\det(\mathcal{G})^{\otimes 2n} \cong (\mathrm{Const}(\chi, D)^{2n})^{\deg}$. Therefore if we take α to be any N'th root of $1/\mathrm{Const}(\chi, D)$, then $\mathcal{G}\otimes(\alpha)^{\deg}$ is pure of weight zero, and has $G_{arith} = G_{geom}$. Then apply Deligne's equidistribution theorem, cf. [Ka–Sar, RMFEM, 9.2.6]. QED

Chapter 8: Average Order of Zero in Twist Families

8.0 The basic setting

(8.0.1) In this section, we work over a finite field k of **odd** characteristic. We give ourselves data (C/k, D, ℓ, r, $\mathcal{F}$, χ, ι, w) as in 7.0. We suppose that after extension of scalars from k to $\overline{k}$, our data (C/k, D, ℓ, r, $\mathcal{F}$, χ) satisfies all the hypotheses of Theorem 5.5.1.

(8.0.2) We further suppose that $\mathcal{F}$(w/2) is symplectically self–dual on C/k, and that χ has order 2. Then, by Poincaré duality, $\mathcal{G}$((w+1)/2) is orthogonally self–dual as a lisse sheaf on

$$X := \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}).$$

By Theorem 5.5.1, $\mathcal{G}$ has G_{geom} either SO(N) or O(N).

8.1 Definitions of three sorts of analytic rank

(8.1.1) Given a finite extension E/k, and f in X(E), we define the **analytic rank** of $\mathcal{G}$ at (E, f), denoted $\mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E, f)$, to be the order of vanishing of

$$\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2))$$

at T = 1, i.e., $\mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E, f)$ is the multiplicity of 1 as generalized eigenvalue of $\mathrm{Frob}_{E,f}$ on $\mathcal{G}$((w+1)/2):

$$\mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E, f) := \mathrm{ord}_{T=1}\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)).$$

(8.1.2) For each n ≥ 1, denote by E_n/E the extension of E of degree n.

(8.1.3) We define the **quadratic analytic rank** of $\mathcal{G}$ at (E, f), denoted $\mathrm{rank}_{\mathrm{quad,\ an}}(\mathcal{G}, E, f)$ to be the sum of the orders of vanishing of

$$\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2))$$

at T = 1 and at T = −1, i.e., $\mathrm{rank}_{\mathrm{quad,\ an}}(\mathcal{G}, E, f)$ is the sum of the multiplicities of 1 and of −1 as generalized eigenvalues of $\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)$. More simply,

$$\mathrm{rank}_{\mathrm{quad,\ an}}(\mathcal{G}, E, f) := \mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E_2, f).$$

(8.1.4) We define the **geometric analytic rank** of $\mathcal{G}$ at (E, f), denoted $\mathrm{rank}_{\mathrm{geom,\ an}}(\mathcal{G}, E, f)$, to be the sum of the orders of vanishing of

$$\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2))$$

at all roots of unity, i.e., $\mathrm{rank}_{\mathrm{geom,\ an}}(\mathcal{G}, f)$ is the sum of the multiplicities of all roots of unity as generalized eigenvalues of $\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)$. More simply,

$$\mathrm{rank}_{\mathrm{geom,\ an}}(\mathcal{G}, E, f) := \lim_{n \to \infty}\mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E_{n!}, f).$$

8.2 Relation to Mordell–Weil rank

(8.2.1) The terminology "analytic rank" is motivated by the Birch and Swinnerton–Dyer conjectures for the ranks of abelian varieties over function fields with finite constant fields. Suppose the sheaf $\mathcal{F}$ arises as the middle extension of the H^1 along the fibres of (the spreading out to some dense open set in C of) an abelian variety A/K, K the function field k(C). For each finite extension E/k and each f in X(E), we form the quadratic twist of A by f, getting an abelian variety $A\otimes\chi_2(f)/EK$. The Birch and Swinnerton–Dyer conjecture for $A\otimes\chi_2(f)/EK$ asserts that its Mordell–Weil rank is given by

$$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK) = \mathrm{rank}_{an}(\mathcal{G}, E, f).$$

This same BSD conjecture, applied now to the same twist but viewed over E_2K, says

$$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K) = \mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f).$$

Because we assume that A/K has a geometrically irreducible $\mathcal{F}$, A/K has no fixed part, even over $\bar{E}K$, and neither does any quadratic twist of it. Therefore $(A\otimes\chi_2(f))(\bar{E}K)$ is a finitely generated group. So writing $\bar{E}K$ as the increasing union of finite constant field extensions $E_{n!}K$ of EK, the BSD conjecure applied to all of these predicts that

$$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/\bar{E}K) = \mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f).$$

(8.2.2) In the function field over a finite field case, we have a priori inequalities

$$0 \le \mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK) \le \mathrm{rank}_{an}(\mathcal{G}, E, f),$$
$$0 \le \mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K) \le \mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f),$$
$$0 \le \mathrm{rank}_{MW}(A\otimes\chi_2(f)/\bar{E}K) \le \mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f).$$

8.3 Theorems on average analytic ranks, and on average Mordell–Weil rank

(8.3.1) Under the hypotheses introduced in 8.0 above, we know that $\mathcal{G}((w+1)/2)$ is orthogonally self–dual, and that G_{geom} is either SO or O. Thus we have

$$SO \subset G_{geom} \subset G_{arith} \subset O.$$

See Proposition 5.5.2 for various conditions which insure that G_{geom} is O(N) rather than SO(N). In particular, recall that G_{geom} is O(N) if N is odd.

(8.3.2) We will consider successively the three possibilities:

$$G_{geom} = G_{arith} = O,$$
$$G_{geom} = G_{arith} = SO,$$
$$G_{geom} = SO,\ G_{arith} = O.$$

Theorem 8.3.3 Hypotheses as in 8.0 above, suppose G_{geom} is the full orthogonal group O. If we take the limit over finite extensions E/k large enough that X(E) is nonempty, we get the following tables of limit formulas. In these tables, the number in the third column is the limit, as $\#E \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X(E)$	$\mathrm{rank}_{an}(\mathcal{G}, E, f)$	1/2,
$X(E)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f)$	1,
$X(E)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f)$	1.

More precisely, for each finite extension E/k, and each value of $\varepsilon = \pm 1$, denote by $X_{\mathrm{sign}\ \varepsilon}(E)$ the subset of $X(E)$ consisting of those points f in $X(E)$ such that

$$\det(-\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)) = \varepsilon.$$

Then we have the following table of limit formulas:

If N is even:

$X_{\mathrm{sign}\ -}(E)$	$\mathrm{rank}_{an}(\mathcal{G}, E, f)$	1,
$X_{\mathrm{sign}\ +}(E)$	$\mathrm{rank}_{an}(\mathcal{G}, E, f)$	0,
$X_{\mathrm{sign}\ -}(E)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f)$	2,
$X_{\mathrm{sign}\ +}(E)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f)$	0,
$X_{\mathrm{sign}\ -}(E)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f)$	2,
$X_{\mathrm{sign}\ +}(E)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f)$	0.

If N is odd:

$X_{\mathrm{sign}\ -}(E)$	$\mathrm{rank}_{an}(\mathcal{G}, E, f)$	1,
$X_{\mathrm{sign}\ +}(E)$	$\mathrm{rank}_{an}(\mathcal{G}, E, f)$	0,
$X_{\mathrm{sign}\ -}(E)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f)$	1,
$X_{\mathrm{sign}\ +}(E)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f)$	1,
$X_{\mathrm{sign}\ -}(E)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f)$	1,
$X_{\mathrm{sign}\ +}(E)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f)$	1.

proof Denote by N the rank of $\mathcal{G}$. The sheaf $\mathcal{G}((w+1)/2)$ is given to us as a lisse $\overline{\mathbb{Q}}_\ell$–sheaf. Any such sheaf is obtained by extension of scalars from a lisse F_λ–sheaf, for F_λ some finite extension of $\mathbb{Q}_\ell$ [Ka–Sar, RMFEM, 9.07]. So each characteristic polynomial

$$\det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2))$$

is a degree N polynomial over F_λ. But F_λ has only finitely many extensions inside $\overline{F}_\lambda$ of degree at most N, so all the reciprocal roots of all these characteristic polynomials all lie in a finite extension L_λ/F_λ. But L_λ contains only finitely many roots of unity, say $M = \#\mu_\infty(L_\lambda)$.

Via the given embedding $\iota : \overline{\mathbb{Q}}_\ell \to \mathbb{C}$, the polynomial

$$\iota\det(1 - \mathrm{TFrob}_{E,f} \mid \mathcal{G}((w+1)/2))$$

is the characteristic polynomial of a unique conjugacy class $\theta(E, f)$ in $O(N, \mathbb{R})$.

Our first task is to define the **reduced characteristic polynomial**

$$\mathrm{Rdet}(1 - T\gamma)$$

for an element γ in $O(N, \mathbb{R})$, cf. [deJ–Ka, 6.7].

If N is even, then every element γ in $O_{\mathrm{sign}\ -}(N, \mathbb{R})$ has both ± 1 as eigenvalues, and we define

$$\mathrm{Rdet}(1 - T\gamma) := \det(1 - T\gamma)/(1 - T^2),\ \gamma \text{ in } O_{\mathrm{sign}\ -}(N, \mathbb{R}),\ N \text{ even}.$$

If N is even and γ lies in $O_{\mathrm{sign}\ +}(N, \mathbb{R})$, we define

$$\mathrm{Rdet}(1 - T\gamma) := \det(1 - T\gamma),\ \gamma \text{ in } O_{\mathrm{sign}\ +}(N, \mathbb{R}),\ N \text{ even}.$$

If N is odd, then every element γ in $O_{\mathrm{sign}\ \varepsilon}(N, \mathbb{R})$ has $-\varepsilon$ as an eigenvalue and we define

$$\mathrm{Rdet}(1 - T\gamma) := \det(1 - T\gamma)/(1 + \varepsilon T),\ \gamma \text{ in } O_{\mathrm{sign}\ \varepsilon}(N, \mathbb{R}),\ N \text{ odd}.$$

The function $\gamma \mapsto \mathrm{Rdet}(1 - T\gamma)$ is a continuous central function on $O(N, \mathbb{R})$ with values in the space of $\mathbb{R}$–polynomials of degree $\le N$.

We denote by Z the closed set of $O(N, \mathbb{R})$ defined by the vanishing of the function

$$\gamma \mapsto \prod_{\xi \text{ in } \mu_M(\mathbb{C})} \mathrm{Rdet}(1 - \gamma\xi).$$

The set Z is visibly invariant by $O(N, \mathbb{R})$–conjugation, and has measure zero for Haar measure, cf. [deJ–Ka, 6.9].

For each γ in $O(N, \mathbb{R})$, and each integer $n \ge 1$, we define

$\mathrm{mult}_n(\gamma)$:= the sum of the multiplicities of all n'th roots of unity as eigenvalues of γ.

The functions mult_1, mult_2, and mult_M are each bounded central $\mathbb{Z}$–valued functions on $O(N, \mathbb{R})$, which are continuous outside of Z. Outside of Z, they agree with the following locally constant functions on $O(N, \mathbb{R})$ [recall that the M in mult_M is the number of roots of unity in L_λ]:

	$O_{\mathrm{sign}\ -}(N, \mathbb{R})$ N even	$O_{\mathrm{sign}\ +}(N, \mathbb{R})$ N even	$O_{\mathrm{sign}\ -}(N, \mathbb{R})$ N odd	$O_{\mathrm{sign}\ +}(N, \mathbb{R})$ N odd
mult_1	1	0	1	0
mult_2	2	0	1	1
mult_M	2	0	1	1

The key point about these multiplicity functions is this: for any finite extension E/k and any point f in X(E), we have

$$\mathrm{rank}_{an}(\mathcal{G}, E, f) = \mathrm{mult}_1(\theta(E, f)),$$

$$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f) = \mathrm{mult}_2(\theta(E, f)),$$

$$\mathrm{rank}_{\mathrm{geom,\ an}}(\mathcal{G}, E, f) = \mathrm{mult}_M(\theta(E, f)).$$

For each finite extension E/k, and each value of $\varepsilon = \pm 1$, denote by $X_{\mathrm{sign}\ \varepsilon}(E)$ the subset of X(E) consisting of those points f in X(E) such that

$$\det(-\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)) = \varepsilon.$$

For each choice of $\varepsilon = \pm 1$, as $\#E \to \infty$,

$$\#X_{\mathrm{sign}\ \varepsilon}(E)/\#X(E) \to 1/2,$$

and the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\mathrm{sign}\ \varepsilon}(E)}$ become equidistributed for the Haar measure of total mass one on the space $O_{\mathrm{sign}\ \varepsilon}(N, \mathbb{R})^{\#}$. There is a standard extension of this result, to more general functions, cf. [Ka–Sar, RMFEM, AD11.4], which will be useful for us below. Let Z be any closed subset of $O(N)_{\mathbb{R}}$ of Haar measure zero which is stable by $O(N)_{\mathbb{R}}$–conjugation, and let g be a bounded, $\mathbb{C}$–valued central function on $O(N)_{\mathbb{R}}$ whose restriction to $O(N)_{\mathbb{R}} - Z$ is continuous. For such a function g we still have the integral formula

$$\int_{O(N,\ \mathbb{R})} g(A)dA = \lim_{\#E \to \infty} (1/\#X(E))\Sigma_{f \text{ in } X(E)}\ g(\theta(E, f)).$$

If we apply this to g×(char function of $O_{\mathrm{sign}\ \varepsilon}(N, \mathbb{R})$), we get the integral formula

$$\int_{O_{\mathrm{sign}\ \varepsilon}(N,\ \mathbb{R})} g(A)dA = \lim_{\#E \to \infty} (1/\#X_{\mathrm{sign}\ \varepsilon}(E))\Sigma_{f \text{ in } X_{\mathrm{sign}\ \varepsilon}(E)}\ g(\theta(E, f)),$$

in which the dA on $O_{\mathrm{sign}\ \varepsilon}(N, \mathbb{R})$ is the restriction of Haar measure, but now normalized to give $O_{\mathrm{sign}\ \varepsilon}(N, \mathbb{R})$ mass one.

We need only take for g successively the functions mult_1, mult_2, and mult_M. Their averages over Frobenii $\theta(E, f)$ are precisely the average analytic ranks in question. Their integrals are easy to compute, since these functions agree, outside a set of measure zero, with the locally constant functions mult_1, mult_2, and mult_M in the table above. QED

Corollary 8.3.4 Hypotheses as in Theorem 8.3.3 above, suppose in addition that the sheaf $\mathcal{F}$ arises as the middle extension of the H^1 along the fibres of (the spreading out to some dense open set in C of) an abelian variety A/K, K the function field k(C). Then we have the following tables of limsup results for the average Mordell–Weil ranks of quadratic twists. In these tables, the number in the third column is an upper bound for the limsup, as $\#E \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column. In those cases where the limsup is 0, the limit exists and is zero, and in those cases we have written "= 0" in the third column.

X(E))	$\mathrm{rank}_{MW}(A \otimes \chi_2(f)/EK)$	$\leq 1/2$,
X(E)	$\mathrm{rank}_{MW}(A \otimes \chi_2(f)/E_2K)$	≤ 1,
X(E)	$\mathrm{rank}_{MW}(A \otimes \chi_2(f)/\bar{E}K)$	≤ 1.

More precisely, for each finite extension E/k, and each value of $\varepsilon = \pm 1$, denote by $X_{\text{sign } \varepsilon}(E)$ the subset of X(E) consisting of those points f in X(E) such that

$$\det(-\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)) = \varepsilon.$$

Then as $\#E \to \infty$, $\#X_{\text{sign } \varepsilon}(E)/\#X(E) \to 1/2$, and we have the following tables:

If N is even:

$X_{\text{sign } -}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK)$	≤ 1,
$X_{\text{sign } +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK)$	$= 0$,
$X_{\text{sign } -}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K)$	≤ 2,
$X_{\text{sign } +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K)$	$= 0$,
$X_{\text{sign } -}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/\bar{E}K)$	≤ 2,
$X_{\text{sign } +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/\bar{E}K)$	$= 0$.

If N is odd:

$X_{\text{sign } -}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK)$	≤ 1,
$X_{\text{sign } +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK)$	$= 0$,
$X_{\text{sign } -}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K)$	≤ 1,
$X_{\text{sign } +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K)$	≤ 1,
$X_{\text{sign } -}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/\bar{E}K)$	≤ 1,
$X_{\text{sign } +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/\bar{E}K)$	≤ 1.

proof Immediate from Theorem 8.3.3 and the a priori inequalities 8.2.2 bounding Mordell–Weil rank by analytic rank. QED

Example 8.3.4.1 Suppose in 8.3.4 we take for A/K an elliptic curve E/K which has multiplicative reduction at some $\bar{k}$–valued point β of C–D. Then $\mathcal{F}$ has unipotent nontrivial monodromy at β. By Proposition 5.5.2, part 1), $\mathcal{G}$ has G_{geom} the full orthogonal group O(N).

(8.3.5) We now turn to the two cases where G_{geom} is SO rather than O. Recall from Proposition 5.5.2 that if G_{geom} is SO, then the rank N of $\mathcal{G}$ is even.

Theorem 8.3.6 Hypotheses as in 8.0 above, suppose $G_{geom} = G_{arith} = SO$. For every finite extension E/k, $X_{sign\ -}(E)$ is empty, and we get the following table of limit formulas. In the table, the number in the third column is the limit, as $\#E \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{sign\ +}(E)$	$\mathrm{rank}_{an}(\mathcal{G}, E, f)$	0,
$X_{sign\ +}(E)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f)$	0,
$X_{sign\ +}(E)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f)$	0.

proof As N is even and G_{arith} is $SO(N) = O_{sign\ +}(N, \mathbb{R})$, all $\theta(E, f)$ lie in and are equidistributed in $O_{sign\ +}(N, \mathbb{R})$, where all three mult functions (introduced in the proof of Theorem 8.3.3) vanish outside Z. QED

Corollary 8.3.7 Hypotheses as in Theorem 8.3.6 above, suppose in addition that the sheaf $\mathcal{F}$ arises as the middle extension of the H^1 along the fibres of (the spreading out to some dense open set in C of) an abelian variety A/K, K the function field k(C). Then we have the following table of limit formulas (same format as in 8.3.6 above) for the average Mordell–Weil ranks of quadratic twists.

$X_{sign\ +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK)$	0,
$X_{sign\ +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K)$	0,
$X_{sign\ +}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/\bar{E}K)$	0.

proof Immediate from Theorem 8.3.6 and the a priori inequalities 8.2.2 bounding Mordell–Weil rank by analytic rank. QED

Theorem 8.3.8 Hypotheses as in 8.0 above, suppose $G_{geom} = SO$ and $G_{arith} = O$. For finite extensions E/k of **even** degree, $X_{sign\ -}(E)$ is empty, and we get the following table of limit formulas over E/k of even degree. In the table, the number in the third column is the limit, as $\#E \to \infty$ over extensions E/k of **even** degree, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{sign\ +}(E)$	$\mathrm{rank}_{an}(\mathcal{G}, E, f)$	0,
$X_{sign\ +}(E)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}, E, f)$	0,
$X_{sign\ +}(E)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}, E, f)$	0.

For finite extensions E/k of **odd** degree, $X_{\mathrm{sign}\,+}(E)$ is empty, and we get the following table of limit formulas over E/k of odd degree. In the table, the number in the third column is the limit, as $\#E \to \infty$ over extensions E/k of **odd** degree, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{\mathrm{sign}\,-}(E)$	$\mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E, f)$	1,
$X_{\mathrm{sign}\,-}(E)$	$\mathrm{rank}_{\mathrm{quad,\,an}}(\mathcal{G}, E, f)$	2,
$X_{\mathrm{sign}\,-}(E)$	$\mathrm{rank}_{\mathrm{geom,\,an}}(\mathcal{G}, E, f)$	2.

proof For E/k of even degree, the $\theta(E, f)$ land in and are equidistributed in $O_{\mathrm{sign}\,+}(N, \mathbb{R})$ where all three mult functions vanish outside Z. For E/k of odd degree, the $\theta(E, f)$ land in and are equidistributed in $O_{\mathrm{sign}\,-}(N, \mathbb{R})$ where the three mult functions are, respectively, the constants 1, 2, 2 outside Z. QED

Corollary 8.3.9 Hypotheses as in Theorem 8.3.8 above, suppose in addition that the sheaf $\mathcal{F}$ arises as the middle extension of the H^1 along the fibres of (the spreading out to some dense open set in C of) an abelian variety A/K, K the function field k(C). Then we have the following results for the Mordell–Weil ranks of quadratic twists.

For finite extensions E/k of **even** degree, $X_{\mathrm{sign}\,-}(E)$ is empty, and we get the following table of limit formulas over E/k of even degree. In the table, the number in the third column is the limit, as $\#E \to \infty$ over extensions E/k of **even** degree, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{\mathrm{sign}\,+}(E)$	$\mathrm{rank}_{\mathrm{MW}}(A\otimes\chi_2(f)/EK)$	0,
$X_{\mathrm{sign}\,+}(E)$	$\mathrm{rank}_{\mathrm{MW}}(A\otimes\chi_2(f)/E_2K)$	0,
$X_{\mathrm{sign}\,+}(E)$	$\mathrm{rank}_{\mathrm{MW}}(A\otimes\chi_2(f)/\bar{E}K)$	0.

For finite extensions E/k of **odd** degree, $X_{\mathrm{sign}\,+}(E)$ is empty, and we get the following table of upper bounds for limsups over E/k of odd degree. In the table, the number in the third column is an upper bound for the limsup, as $\#E \to \infty$ over extensions E/k of **odd** degree, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{\mathrm{sign}\,-}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/EK)$	≤ 1,
$X_{\mathrm{sign}\,-}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/E_2K)$	≤ 2,
$X_{\mathrm{sign}\,-}(E)$	$\mathrm{rank}_{MW}(A\otimes\chi_2(f)/\ltimes EK)$	≤ 2.

8.4 Examples of input $\mathcal{F}$'s with small G_{geom}

(8.4.1) We wish to give examples of abelian schemes $p : \mathcal{A} \to U$, U a dense open set in C, such that the middle extension $\mathcal{F}$ of $R^1p_*\overline{\mathbb{Q}}_\ell$ is geometrically irreducible. The simplest way to do this is to exhibit families of curves $\pi : \mathcal{Y} \to U$ whose $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is not only geometrically irreducible, but has G_{geom} the full symplectic group Sp(2d). One then takes for $p : \mathcal{A} \to U$ the family of Jacobians. In this case $R^1\pi_*\overline{\mathbb{Q}}_\ell = R^1p_*\overline{\mathbb{Q}}_\ell$. We refer to [Ka–Sar, RMFEM, Chapter 10] for a plethora of examples of such families of curves. [In those examples, the base is an open set V in $\mathbb{P}^1$. After any nonconstant map f: $C \to \mathbb{P}^1$, the pullback family over $f^{-1}(V)$ still has G_{geom} = Sp(2d), simply because Sp(2d) is connected.] In fact, in most "natural" examples where we know that $\mathcal{F}$ is geometrically irreducible, we know it because we can show G_{geom} is Sp(2d).

(8.4.2) However, there is a general procedure to construct, for every integer $d \ge 2$, examples of d–dimensional abelian varieties A/K for whose $\mathcal{F}$ the group G_{geom} is a quite small irreducible subgroup of Sp(2d). Begin with a dense open set U in C, and an elliptic curve $\mathcal{E}$/U whose j–invariant is nonconstant. Given an integer $d \ge 2$, pick a finite subgroup Γ of the orthogonal group $O(d, \mathbb{Z})$ such that Γ acts irreducibly on $\mathbb{C}^d$. [For instance, we might take Γ to be the symmetric group S_{d+1} in its augmentation representation.] Pick an integer N such that the maximal prime–to–p quotient of Γ is generated by N elements. Shrink U if necessary, so that $(C-U)(\overline{k})$ consists of at least N+1 points. **Suppose** there exists a finite etale Γ–torsor

$$V \to U$$

such that V/k is geometrically connected. Then we take the abelian scheme $\mathcal{E}^d$/U, think of $\mathcal{E}^d$ as $\mathcal{E}\otimes_{\mathbb{Z}}\mathbb{Z}^d$, and twist it by the covering V/U, having Γ act on $\mathcal{E}\otimes_{\mathbb{Z}}\mathbb{Z}^d$ as $(\mathrm{id}_E)\otimes$(given rep. of Γ on $\mathbb{Z}^d$). This twisted abelian scheme is a d–dimensional $\mathcal{A}$/U. Its $\mathcal{F}(\mathcal{A}/U)$ is canonically a tensor product

$$\mathcal{F}(\mathcal{A}/U) = \mathcal{F}(\mathcal{E}/U)\otimes((\overline{\mathbb{Q}}_\ell)^d \text{ as } \Gamma\text{–representation}).$$

In terms of this decomposition, G_{geom} is the irreducible subgroup $SL(2)\otimes\Gamma$ of Sp(2d). This follows from a form of Goursat's Lemma, cf. 9.7.3, and the fact that $\mathcal{F}(\mathcal{E}/U)$ has G_{geom} the **connected** group SL(2).

(8.4.3) Can we construct a finite etale Γ–torsor

$$V \to U$$

such that V/k is geometrically connected? The answer is yes, **if** we allow ourselves a finite extension of the constant field. By the positive solution to the Abhyankar Conjecture [Harb–AC], we know that Γ is a quotient of $\pi_1(U\otimes_k\bar{k})$, i.e., there exists a connected finite etale galois Γ–torsor

$$V \to U\otimes_k\bar{k}.$$

Since $\bar{k}$ is the union of finite extensions of k, for some finite extension k_1 of k, this diagram descends to a connected Γ–torsor

$$V_1 \to U\otimes_k k_1.$$

Thus we get a d–dimensional $\mathcal{A}/U\otimes_k k_1$ whose $\mathcal{F} | U\otimes_k k_1$ has G_{geom} the irreducible subgroup $SL(2)\otimes(\Gamma$ acting on $(\bar{\mathbb{Q}}_\ell)^d)$ of Sp(2d).

(8.4.4) We do not know if we can avoid the necessity of making a finite constant field extension k_1/k in general. But there are some elementary cases where no constant field extension is necessary. Here is one such example.

(8.4.5) Suppose that the characteristic p does not divide d(d+1). Then the finite flat map

$$f : \mathbb{A}^1 \to \mathbb{A}^1$$
$$f : X \mapsto (-1/d)(X^{d+1} - (d+1)X)$$

is weakly supermorse, cf. [Ka–ACT, 5.5.2]. [This means that the d+1 = deg(f) is prime to p, that the differential df has d distinct zeroes, and that f separates these zeroes. Here the zeroes of df are the d'th roots of unity, and $f(\xi) = \xi$ for ξ any d'th root of unity.] The map f makes $\mathbb{A}^1 - f^{-1}(\mu_d)$ a finite etale covering of $\mathbb{A}^1 - \mu_d$ of degree d+1. The lisse sheaf $\mathcal{F}$ on the base $\mathbb{A}^1 - \mu_d$ defined as

$$\mathcal{F} := \text{Kernel of Trace} : f_*\bar{\mathbb{Q}}_\ell \to \bar{\mathbb{Q}}_\ell$$

is then an irreducible tame reflection sheaf, whose G_{geom} is the full symmetric group S_{d+1}, cf. [Ka–ACT, 5.5.3.6] and [Ka–ESDE, proof of 7.10.2.3]. [In more down to earth terms, over $\mathbb{F}_p(T)$, the equation

$$(-1/d)(X^{d+1} - (d+1)X) = T$$

has galois group S_{d+1}, and keeps this same galois group over $\bar{\mathbb{F}}_p(T)$.]

Thus we get an S_{d+1}–torsor

$$V \to \mathbb{A}^1 - \mu_d$$

with $V/\mathbb{F}_p$ geometrically connected.

(8.4.6) Now pick a prime number $\ell_1 > \mathrm{Max}(2g, d+1)$. At the expense of shrinking U, we may assume that C–U contains a closed point $\mathcal{P}$ of degree ℓ_1. Take a nonconstant function g in $L(\mathcal{P})$ (possible by Riemann–Roch). Then g has a simple pole at $\mathcal{P}$ and no other poles. So it defines a finite flat generically etale map of C to $\mathbb{P}^1\otimes_{\mathbb{F}_p}k$ of degree ℓ_1. At the expense of further shrinking U, we may assume that g maps U to $(\mathbb{A}^1 - \mu_d)\otimes_{\mathbb{F}_p}k$. Since ℓ_1 is prime to (d+1)!, a linear disjointness argument shows that the pullback by

$$g : U \to \mathbb{A}^1 - \mu_d$$

of the S_{d+1}–torsor

$$V\otimes_{\mathbb{F}_p}k \to (\mathbb{A}^1 - \mu_d)\otimes_{\mathbb{F}_p}k$$

is an S_{d+1}–torsor

$$g^*V \to U$$

whose total space remains geometrically connected.

(8.4.7) There is another way to construct abelian schemes

$$p : \mathcal{A} \to U$$

of any dimension $d \geq 2$ over open sets U of $\mathbb{P}^1$ such that the middle extension $\mathcal{F}$ of $R^1p_*\overline{\mathbb{Q}}_\ell$ is geometrically irreducible, but whose G_{geom} is a quite small irreducible subgroup of Sp(2d) (though not as small as in the previous construction). We start with a dense open set V in $\mathbb{P}^1$, and an elliptic curve $\pi : \mathcal{E} \to V$ whose j–invariant is nonconstant. We form the (geometrically irreducible, because j is nonconstant) middle extension $\mathcal{F}_1$ of $R^1\pi_*\overline{\mathbb{Q}}_\ell$. Again because $\mathcal{E}/V$ has nonconstant j–invariant, $\mathcal{E}/V$ has bad reduction at some point of $\mathbb{P}^1 - V$. By the Néron–Ogg–Shafarevich criterion of good reduction, the middle extension $\mathcal{F}_1$ is not everywhere lisse on $\mathbb{P}^1$, i.e., $\mathrm{Sing}(\mathcal{F}_1)$ is nonempty. At the expense of extending the ground field k, we may assume that $\mathrm{Sing}(\mathcal{F}_1)$ contains a k–rational point, and that $\mathbb{P}^1 - \mathrm{Sing}(\mathcal{F}_1)$ contains at least d–1 k–rational points. Pick one point P_1 in $\mathrm{Sing}(\mathcal{F}_1)(k)$, and pick d–1 distinct k–rational points $P_2, ..., P_d$ in $\mathbb{P}^1 - \mathrm{Sing}(\mathcal{F}_1)$. Pick a coordinate x for the source $\mathbb{P}^1$ such that none of the P_i is ∞. Then consider the function

$$f(x) := 1/\textstyle\prod_i(x - x(P_i)).$$

This function is a finite flat map of degree d from $\mathbb{P}^1$ to itself, which is finite etale over ∞ in the target (and hence finite etale over some dense open set of the target). In the fibre $f^{-1}(\infty)$, there is precisely one point, namely P_1, in $\mathrm{Sing}(\mathcal{F}_1)$. So by the Irreducible Induction Criterion 3.3.1, the direct image $f_*\mathcal{F}_1$ on the target $\mathbb{P}^1$ is a geometrically irreducible middle extension, of generic rank 2d. This sheaf $f_*\mathcal{F}_1$ is precisely the middle extension sheaf $\mathcal{F}$ attached to a spreading out of a certain d–dimensional abelian variety A over the function field k(t) of the target $\mathbb{P}^1$. Namely, denote by E/k(x) the generic fibre of the elliptic curve $\mathcal{E}/V$ we started with. Then the A in question is the Weil restriction of scalars, from k(x) to k(t), t := f(x), of E.

$$A := R_{k(x)/k(t)}(E).$$

Over the galois closure $k(x)^{gal}/k(t)$ of the separable extension k(x)/k(t), A becomes the product of the d conjugates of E/k(x) by the d embeddings of k(x) into $k(x)^{gal}$ which are the identity on k(t). This means that for (the spreading out of) our A/k(t), the connected component $(G_{geom})^0$ lies in the d−fold product of SL(2) with itself.

8.5 Criteria for when G_{geom} is SO rather than O

(8.5.1) This section is a complement to Proposition 5.5.2 and to the discussion in section 7.4. We continue to work over a finite field k of **odd** characteristic. We fix data

$$(C/k, D, \ell, r, \mathcal{F}, \chi_2, \iota, w)$$

as in 8.0.1−2. We also fix a choice α_k of Sqrt(q) in $\overline{\mathbb{Q}}_\ell$, and agree to use powers of this α_k in forming Tate twists by half−integers. Thus r is even, and $\mathcal{F}(w/2)$ is symplectically self−dual and ι−pure of weight zero.

(8.5.2) We now make two further assumptions.

(8.5.2.1) $\mathcal{F}$ is everywhere tamely ramified.

(8.5.2.2) The degree d of the divisor D satisfies

$$d \geq 4g+4, \text{ and we have}$$
$$2g - 2 + d > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}), 72r).$$

The first assumption, that $\mathcal{F}$ is everywhere tame, is essential. The second assures us that Theorem 5.5.1 applies, whatever the effective divisor D.

(8.5.3) We form the sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi_2,C,D}(\mathcal{F}).$$

We know that $\mathcal{G}((w+1)/2)$ is orthogonally self−dual as a lisse sheaf on

$$X := \mathit{Fct}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}}).$$

(8.5.4) By Theorem 5.5.1, $\mathcal{G}$ has G_{geom} either SO(N) or O(N), N being rank($\mathcal{G}$). We wish to give some more criteria to decide which of these two cases we are in. The idea is very simple. As explained in section 7.4, we can numerically decide this question by computing the determinants of Frobenii acting on various stalks $\mathcal{G}_f((w+1)/2)$ of $\mathcal{G}((w+1)/2)$ over various extension fields, and seeing how their signs vary. For any finite extension E/k, and any f in X(E), the stalk $\mathcal{G}_f((w+1)/2)$ is the cohomology group

$$\mathcal{G}_f((w+1)/2) := H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}))((w+1)/2),$$

and the action of $\mathrm{Frob}_{E,f}$ on $\mathcal{G}_f$ is the action of Frob_E on this cohomology group. As explained in 7.0.6.4, this leads to

$$L(C\otimes_k E, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})((w+1)/2))(T)$$
$$= \det(1 - T\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})((w+1)/2)))$$
$$= \det(1 - T\mathrm{Frob}_{E,f} \mid \mathcal{G}_f((w+1)/2)).$$

Thus $\det(-\mathrm{Frob}_{E,f} \mid \mathcal{G}_f((w+1)/2))$ is the **sign in the functional equation** of the L−function

$$L(C\otimes_k E, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})((w+1)/2))(T).$$

Equivalently, the constant

(8.5.4.1) $$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}) := 1/\det(-\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}))$$

is the product of the sign in the functional equation times an integral power of $\alpha_E := \alpha_k^{\deg(E/k)}$.

(8.5.5) In principle, we can use the theory of local constants ([De–Const], [Lau–TFC]) to compute this sign, or more precisely to see whether or not it varies with f in a fixed X(E). In practice, this is not so easy to carry out, and that is why in Theorem 8.5.7 below the hypotheses are somewhat restrictive.

(8.5.6) Recall from 7.4 that we have $G_{geom} = O(N)$ if and only if the sign varies as f runs over X(E) for any (or for every) sufficiently large finite extension E/k. We have $G_{geom} = SO = G_{arith}$ (for $\mathcal{G}((w+1)/2)$) if and only if the constant is always +1 for every f over every finite extension. And we have $G_{geom} = SO$ but $G_{arith} = O$ if and only if the constant is equal to $(-1)^{\deg(E/k)}$ for every f in every X(E).

Theorem 8.5.7 Hypotheses as in 8.5.1 and 8.5.2 above, suppose in addition that each point of $\mathrm{Sing}(\mathcal{F})$ occurs in D with even (possibly zero) multiplicity. Then we have the following results.

1) If at every geometric point β of $\mathrm{Sing}(\mathcal{F})$, $\dim(\mathcal{F}/\mathcal{F}^{I(\beta)})$ is even, then $\mathcal{G}((w+1)/2))$ has

$$G_{geom} = SO.$$

Moreover, we have

G_{arith} is SO if $\varepsilon(k, \mathcal{F})$ is an integral power of α_k,

G_{arith} is O if $\varepsilon(k, \mathcal{F})$ is $(-1)\times$(an integral power of α_k).

2) If there exists a geometric point β of $\mathrm{Sing}(\mathcal{F})$ for which $\dim(\mathcal{F}/\mathcal{F}^{I(\beta)})$ is odd, then $\mathcal{G}((w+1)/2))$ has $G_{geom} = O = G_{arith}$.

proof The key point is this: for f in any X(E), $\mathcal{L}_{\chi_2(f)}$ and $\mathcal{F}$ have **disjoint ramification** on $C\otimes_k E$. This disjointness allows us to apply Deligne's formula [De–Const, 9.5] (valid without assuming $\mathcal{F}$ part of a compatible system, thanks to Laumon [Lau–TFC, 3.2.1.1]) to compute the ratio of signs

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2))/\varepsilon(E, \mathcal{F}((w+1)/2)).$$

To carry this out, extend scalars from k to E, and work over E. Denote by EK the function field of $C\otimes_k E$. At each closed point x of $C\otimes_k E$, we denote by $\mathcal{F}(x)$ and $\mathcal{L}_{\chi_{2,E}(f)}(x)$ the representations of the decomposition group D(x) given by $\mathcal{F}$ and by $\mathcal{L}_{\chi_2(f)}$ respectively. We also pick a uniformizing parameter π_x at x. We use local class field theory to view continuous $(\bar{\mathbb{Q}}_\ell)^\times$–valued characters of D_x as characters of $EK_x^\times$, where EK_x denotes the x–adic completion of EK.

Because $\mathcal{F}$ is everywhere tame, its Artin conductor $a_x(\mathcal{F})$ at x is just its drop as a representation of the inertia group I(x):

$$a_x(\mathcal{F}) = \dim(\mathcal{F}(x)/\mathcal{F}(x)^{I(x)}).$$

Because $\mathcal{L}_{\chi_2(f)}$ is everywhere tame, and χ_2 is the quadratic character, we have

$$a_x(\mathcal{L}_{\chi_2(f)}) = 0 \text{ if } \mathrm{ord}_x(f) \text{ is even},$$

$$a_x(\mathcal{L}_{\chi_2(f)}) = 1 \text{ if } \mathrm{ord}_x(f) \text{ is odd}.$$

Deligne's formula [De–Const, 9.5] is

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})/\varepsilon(E, \mathcal{F})$$

$$= [\varepsilon(E, \mathcal{L}_{\chi_2(f)})/\varepsilon(E, \overline{\mathbb{Q}}_\ell)]^r \times [\prod_{x \text{ in } \mathrm{Sing}(\mathcal{F})} \mathcal{L}_{\chi_2(f)}(x)((\pi_x)^{a_x(\mathcal{F})})]$$

$$\times [\prod_{x \text{ in } \mathrm{Sing}(\mathcal{L}_{\chi_2(f)})} (\det\mathcal{F}(x))((\pi_x)^{a_x(\mathcal{L}_{\chi_2(f)})})].$$

Let us make several observations. The first is that since we are trying to track the variation of the sign, and no power of α_E is a nontrivial root of unity, we may work in the quotient group of $(\overline{\mathbb{Q}}_\ell)^\times$ by the multiplicative subgroup generated by α_E. We will write $a \approx b$ if a/b is an integral power of α_E. We have

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}) \approx \varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2)),$$

$$\varepsilon(E, \mathcal{F}) \approx \varepsilon(E, \mathcal{F}((w+1)/2)).$$

The second is that since $\mathcal{F}(w/2)$ is symplectically self–dual, $\det(\mathcal{F}(w/2))$ is trivial, or, equivalently, $\det(\mathcal{F}) = \overline{\mathbb{Q}}_\ell(-wr/2)$. So for every closed point x, every value of $\det\mathcal{F}(x)$ as character of $K_x^\times$ is an integer power of $(\#k(x))^{wr/2}$, and hence an integer power of α_E. So we can throw away the last product if we work modulo powers of α_E, and we find

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})/\varepsilon(E, \mathcal{F})$$

$$\approx [\varepsilon(E, \mathcal{L}_{\chi_2(f)})/\varepsilon(E, \overline{\mathbb{Q}}_\ell)]^r \times [\prod_{x \text{ in } \mathrm{Sing}(\mathcal{F})} \mathcal{L}_{\chi_{2,}(f)}(x)((\pi_x)^{a_x(\mathcal{F})})].$$

Next we observe that because $\mathcal{L}_{\chi_{2,E}(f)}(x)$ is a character of order two, the terms indexed by a point x in $\mathrm{Sing}(\mathcal{F})$ with $a_x(\mathcal{F})$ even are all identically 1, and the terms with $a_x(\mathcal{F})$ odd don't change if in each we replace $a_x(\mathcal{F})$ by 1. Finally, we observe that both the sheaves $\overline{\mathbb{Q}}_\ell$ and $\mathcal{L}_{\chi_2(f)}$, or more precisely their middle extensions from dense opens where they are lisse, are orthogonally self–dual on $C\otimes_k E$. So by Poincaré duality, both of the cohomology groups

$$H^1(C\otimes_k\bar{k}, j_*\mathcal{L}_{\chi_2(f)})(1/2) \text{ and } H^1(C\otimes_k\bar{k}, \overline{\mathbb{Q}}_\ell)(1/2)$$

are symplectically self–dual. Therefore we have

$$\varepsilon(E, j_*\mathcal{L}_{\chi_2(f)}(1/2)) = \varepsilon(E, \overline{\mathbb{Q}}_\ell(1/2)) = 1.$$

Therefore we have

$$\varepsilon(E, j_*\mathcal{L}_{\chi_2(f)}) \approx \varepsilon(E, j_*\mathcal{L}_{\chi_2(f)}(1/2)) = 1,$$
$$\varepsilon(E, \overline{\mathbb{Q}}_\ell) \approx \varepsilon(E, \overline{\mathbb{Q}}_\ell(1/2)) = 1.$$

So we find the following $\approx$ formula for our ratio of constants.

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})/\varepsilon(E, \mathcal{F})$$
$$\approx\prod_{x \text{ in } \mathrm{Sing}(\mathcal{F}) \text{ with } a_x(\mathcal{F}) \text{ odd}} \mathcal{L}_{\chi_2(f)}(x)((\pi_x)^{a_x(\mathcal{F})}).$$
$$\approx\prod_{x \text{ in } \mathrm{Sing}(\mathcal{F}) \text{ with } a_x(\mathcal{F}) \text{ odd}} \mathcal{L}_{\chi_2(f)}(x)(\pi_x).$$

With this formula in hand, we can proceed in a straightforward way. Suppose first there are **no** points where the drop $a_x(\mathcal{F})$ is odd. Then the formula gives

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}) \approx \varepsilon(E, \mathcal{F}),$$

or, equivalently, an equality of signs

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2)) = \varepsilon(E, \mathcal{F}((w+1)/2)).$$

So for each given finite extension E/k, the sign does not vary as f varies in X(E). This lack of variation implies that $\mathcal{G}((w+1)/2)$ has its G_{geom} equal to SO. To determine whether its G_{arith} is O or SO, we must see if the common sign for all f in X(E) depends on the degree of E/k, or not. To do this, we may replace k by any extension of itself of odd degree, and this allows us to assume that X(k) is nonempty. So we pick an f in X(k). We already know (5.5.2) that if G_{geom} is SO, then $\mathcal{G}$ has even rank. So we have

$$\varepsilon(E, \mathcal{F}((w+1)/2))$$
$$= \varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2))$$
$$:= \det(-\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2)))$$
$$= \det(\mathrm{Frob}_E \mid H^1(C\otimes_k\bar{k}, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2)))$$
$$= \det((\mathrm{Frob}_k)^{\deg(E/k)} \mid H^1(C\otimes_k\bar{k}, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2)))$$
$$= \det(\mathrm{Frob}_k \mid H^1(C\otimes_k\bar{k}, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2)))^{\deg(E/k)}$$
$$= \det(-\mathrm{Frob}_k \mid H^1(C\otimes_k\bar{k}, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2)))^{\deg(E/k)}$$
$$= \varepsilon(k, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)}((w+1)/2))^{\deg(E/k)}$$
$$= \varepsilon(k, \mathcal{F}((w+1)/2))^{\deg(E/k)}.$$

Since $\varepsilon(k, \mathcal{F}((w+1)/2))$ is ±1, and is $\approx \varepsilon(k, \mathcal{F})$, we see that the sign varies as $(-1)^{\deg(E/k)}$ if and only if $\varepsilon(k, \mathcal{F}) \approx -1$. This completes the proof of 1).

In order to prove 2), it suffices to find a single finite extension E/k such that as f varies over X(E), the sign changes. So we may extend scalars and reduce to the case where all the points in $\mathrm{Sing}(\mathcal{F})$ are k–rational. At each of them, we pick a uniformizing parameter π_x. Our starting point is the basic formula derived above: for E/k a finite extension, and f in X(E), we have

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})/\varepsilon(E, \mathcal{F}) \approx \prod_{x \text{ in } \mathrm{Sing}(\mathcal{F}) \text{ with } a_x(\mathcal{F}) \text{ odd}} \mathcal{L}_{\chi_2(f)}(x)(\pi_x).$$

But now we are assuming that there are points x in $\mathrm{Sing}(\mathcal{F})$ with $a_x(\mathcal{F})$ odd. At each point x in $\mathrm{Sing}(\mathcal{F})$, the ratio $f/(\pi_x)^{\mathrm{ord}_x(f)}$ is a unit in EK_x which mod squares of units is independent of the auxiliary choice of uniformizing parameter. [This holds because $\mathrm{ord}_x(f)$ is even at each point x in $\mathrm{Sing}(\mathcal{F})$. Indeed, if x lies in D, then f has an even order pole at x, and if x is in $\mathrm{Sing}(\mathcal{F})\cap(C-D)$, then f is a unit at x.] In terms of this unit $f/(\pi_x)^{\mathrm{ord}_x(f)}$, we have the tautological but key identity

$$\mathcal{L}_{\chi_2(f)}(x)(\pi_x) = \chi_2\circ\mathrm{Norm}_{E/k}(\text{the value in } E^\times \text{ of } f/(\pi_x)^{\mathrm{ord}_x(f)} \text{ at } x).$$

To achieve some economy of notation, for x in $\mathrm{Sing}(\mathcal{F})$ and any f in $L(D)\otimes_k E$, we define

$$\bar{f}(x) := \text{the value in E of } f/(\pi_x)^{\mathrm{ord}_x(f)} \text{ at } x.$$

Thus for f in X(E) we have

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})/\varepsilon(E, \mathcal{F}) \approx \prod_{x \text{ in } \mathrm{Sing}(\mathcal{F}) \text{ with } a_x(\mathcal{F}) \text{ odd}} \chi_{2,E}(\bar{f}(x)).$$

For fixed x in $\mathrm{Sing}(\mathcal{F})$, the map

$$\mathrm{Lin}_x: L(D) \to k,$$
$$f \mapsto \bar{f}(x),$$

is a linear form on the k−vector space L(D), and its formation commutes with extension of ground field k. Now X as variety over k is a dense open set in **L(D)**, the affine variety over k whose E−valued points are $L(D)\otimes_k E$ for every E/k. Each of the linear forms Lin_x is an invertible function on the open set X, as is their product

$$\prod := \prod_{x \text{ in } \mathrm{Sing}(\mathcal{F}) \text{ with } a_x(\mathcal{F}) \text{ odd}} \mathrm{Lin}_x.$$

So we may form the lisse, rank one Kummer sheaf $\mathcal{L}_{\chi_2(\prod)}$ on X. In terms of this Kummer sheaf, we have, for every finite extension E/k and every f in X(E),

$$\varepsilon(E, \mathcal{F}\otimes\mathcal{L}_{\chi_2(f)})/\varepsilon(E, \mathcal{F}) \approx \mathrm{Frob}_{E,f} \mid \mathcal{L}_{\chi_2(\prod)}.$$

So the sign varies as f varies over X(E) for large E if and only if $\mathcal{L}_{\chi_2(\prod)}$ is not geometrically constant on X. Now $\mathcal{L}_{\chi_2(\prod)}$ is geometrically constant on X if and only if on $X\otimes_k\bar{k}$, the function $\prod$ is the **square** of another function. The function $\prod$ is the restriction to X of the function $\prod$ on **L(D)**. Since $X\otimes_k\bar{k}$ is open dense in the normal connected $\bar{k}$−scheme $\mathbf{L(D)}\otimes_k\bar{k}$, if $\prod = F^2$ for some function F on $X\otimes_k\bar{k}$, or even for some F in the function field of $X\otimes_k\bar{k}$, that function F must lie in the coordinate ring of $\mathbf{L(D)}\otimes_k\bar{k}$.

To see this, we argue as follows. The coordinate ring of $\mathbf{L(D)}\otimes_k\bar{k}$ is a polynomial ring R in several variables over $\bar{k}$, and $\prod$ is the product of several nonzero linear forms in R. Now R is a U.F.D., and each nonzero linear form is an irreducible element of R. To show that their product is not a square, it suffices to show that the linear forms Lin_x for two different points x are not

$R^\times = \bar{k}^\times$–multiples of each other. Then our $\prod$ is a product of distinct mod $R^\times$ irreducibles, and so by unique factorization it is not a square in R.

But D has large degree, so is very ample. Concretely, it embeds $C(\bar{k})$ into the set of hyperplanes in $L(D)\otimes_k\bar{k}$, by the map

x in $C(\bar{k}) \mapsto$ the hyperplane in $H^0(C\otimes_k\bar{k}, \mathcal{O}(-D))$ consisting of the sections which vanish at x.

And for x in $\mathrm{Sing}(\mathcal{F})$, its hyperplane is precisely the kernel of the linear form Lin_x. Therefore the various linear forms Lin_x for x in $\mathrm{Sing}(\mathcal{F})$ are all distinct mod $R^\times$ irreducibles. Therefore no nonempty partial product of them is a square. This shows that $\mathcal{L}_{\chi_2(\prod)}$ is not geometrically constant on $X\otimes_k\bar{k}$, and completes the proof. QED

8.6 An interesting example

(8.6.1) Let k be a finite field of odd characteristic, ℓ a prime number invertible in k. Over the rational function field $K = k(\lambda)$, we begin with the Legendre curve

$$y^2 = x(x-1)(x-\lambda).$$

Then we form its quadratic twist by $\lambda(\lambda-1)$. This is the curve

$$y^2 = \lambda(\lambda-1)x(x-1)(x-\lambda),$$

which we will name E/K in the following discussion. This curve has good reduction outside of $\{0, 1, \infty\}$, We denote by

$$\pi : \mathcal{E} \to \mathbb{P}^1 - \{0,1,\infty\}$$

the resulting elliptic curve over $\mathbb{P}^1 - \{0,1,\infty\}$.

(8.6.2) Recall that, denoting by

$$j : \mathbb{P}^1 - \{0,1,\infty\} \to \mathbb{P}^1$$

the inclusion, we formed $R^1\pi_*\bar{\mathbb{Q}}_\ell$ on $\mathbb{P}^1 - \{0,1,\infty\}$, and defined

$$\mathcal{F} := j_*R^1\pi_*\bar{\mathbb{Q}}_\ell.$$

The local monodromy of $\mathcal{F} | \mathbb{P}^1 - \{0,1,\infty\}$ is $\mathcal{L}_{\chi_2}\otimes\mathrm{Unip}(2)$ at each of 0, 1, ∞. However, it will be convenient in what follows to pay closer attention to questions of ℓ–adic rationality. With this in mind, we define

$$\mathcal{F}_\ell := j_*R^1\pi_*\mathbb{Q}_\ell.$$

Thus $\mathcal{F}_\ell$ is the natural $\mathbb{Q}_\ell$–form of the $\bar{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ we have been dealing with throughout.

(8.6.3) For each **even** integer $d \geq 144$, we define a divisor D_d in $\mathbb{P}^1$ by $D_d := d\infty$, and form the sheaf

$$\mathcal{G}_d := \mathrm{Twist}_{\chi_2,\mathbb{P}^1,D_d}(\mathcal{F})$$

on the space

$$X_d := Fct(\mathbb{P}^1, D, d, \{0,1\})$$

of degree d polynomials in λ with invertible discriminant and which are invertible at both 0 and 1. The Tate–twisted sheaf $\mathcal{G}_d(1)$ is orthogonally self–dual. According to Theorem 8.5.7, part 1), the group G_{geom} for $\mathcal{G}_d(1)$ is SO(2d). Moreover, the group G_{arith} for $\mathcal{G}_d(1)$ is SO(2d) if the sign in the functional equation for the L–function of E/K is +1, and it is O(2d) if this sign is −1.

(8.6.4) So for each odd prime p, it is natural to ask: what is the sign ε in the functional equation of the L–function of $E/\mathbb{F}_p(\lambda)$, for E the curve

$$y^2 = \lambda(\lambda-1)x(x-1)(x-\lambda)?$$

In terms of the sheaf $\mathcal{F}_\ell$ on $\mathbb{P}^1/\mathbb{F}_p$, this sign is

$$\det(-\mathrm{Frob}_p \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell)(1))$$
$$= \det(-\mathrm{Frob}_p \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell))/p^{\mathrm{rank}(\mathcal{F}_\ell)}.$$

Theorem 8.6.5 The sign $\varepsilon(p) := \det(-\mathrm{Frob}_p \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell)(1))$ is given by

$$\varepsilon(p) = 1 \text{ if } p \equiv 1 \bmod 4,$$
$$\varepsilon(p) = -1 \text{ if } p \equiv 3 \bmod 4.$$

(8.6.6) Before giving the proof, in 8.7 below, we give the main application.

Corollary 8.6.7 In the situation of 8.6.3, fix an odd prime p, and consider for each **even** integer $d \geq 144$ the divisor $D_d := d\infty$, and the sheaf

$$\mathcal{G}_d := \mathrm{Twist}_{\chi_2,\mathbb{P}^1,D_d}(\mathcal{F})$$

on the space

$$X_d := Fct(\mathbb{P}^1, D, d, \{0,1))/\mathbb{F}_p.$$

The Tate–twisted sheaf $\mathcal{G}_d(1)$ is orthogonally self–dual, with group G_{geom}= SO(2d). The group G_{arith} for $\mathcal{G}_d(1)$ is SO(2d) if $p \equiv 1 \bmod 4$, and it is O(2d) if $p \equiv 3 \bmod 4$.

8.7 Proof of Theorem 8.6.5

(8.7.1) The sheaf $\mathcal{F}_\ell$ is everywhere tame on $\mathbb{P}^1$. On $\mathbb{P}^1 - \{0,1,\infty\}$ it is lisse of rank 2, and its stalk vanishes at each of $0,1,\infty$. So the Euler–Poincaré formula gives

$$\chi(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell) = \chi((\mathbb{P}^1 - \{0,1,\infty\})\otimes\overline{\mathbb{F}}_p)\times\mathrm{rank}(\mathcal{F}) = -2.$$

Because $\mathcal{F}_\ell$ is an irreducible middle extension of generic rank > 1, the groups $H^0(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell)$ and $H^2(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell)$ vanish, so we find that

$$\dim H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell) = 2.$$

(8.7.2) Applying the Lefschetz trace formula to $\mathcal{F}_\ell$ on $\mathbb{P}^1$, we have, for any finite extension $E/\mathbb{F}_p$,

$$-\mathrm{Trace}(\mathrm{Frob}_E \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell)) = \sum_{\alpha \text{ in } \mathbb{P}^1(E)} \mathrm{Trace}(\mathrm{Frob}_{E,\lambda} \mid \mathcal{F}_{\ell,\alpha}).$$

The stalk $\mathcal{F}_{\ell,\alpha}$ vanishes at 0, 1, ∞. At any other α in $\mathbb{P}^1(E)$, $\mathcal{E}_\alpha$ is an elliptic curve over E, and $\mathrm{Trace}(\mathrm{Frob}_{E,\lambda} \mid \mathcal{F}_\lambda)$ is

$$\mathrm{Trace}(\mathrm{Frob}_{E,\lambda} \mid \mathcal{F}_\lambda) = \#E + 1 - \#\mathcal{E}_\alpha(E)$$
$$= \#E - \#\{(x,y) \text{ in } E^2 \text{ with } y^2 = \alpha(\alpha-1)x(x-1)(x-\alpha)\}$$
$$= -\sum_{x \text{ in } E} \chi_{2,E}(\alpha(\alpha-1)x(x-1)(x-\alpha)),$$

where we have written $\chi_{2,E}$ for the quadratic character of $E^\times$.

Thus we find

$$\mathrm{Trace}(\mathrm{Frob}_E \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F})) = -\sum_{\alpha \neq 0,1 \text{ in } E} \mathrm{Trace}(\mathrm{Frob}_{E,\alpha} \mid \mathcal{F}_{\ell,\alpha})$$
$$= \sum_{\alpha \neq 0,1 \text{ in } E} \sum_{x \text{ in } E} \chi_{2,E}(\alpha(\alpha-1)x(x-1)(x-\alpha)).$$

But with the usual convention that $\chi_{2,E}(0) = 0$, we can rewrite this as

$$\mathrm{Trace}(\mathrm{Frob}_E \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}))$$
$$= \sum_{\alpha,x \text{ in } E} \chi_{2,E}(\alpha(\alpha-1)x(x-1)(x-\alpha)).$$

(8.7.3) In order to see more clearly what is going on here, we will give the more neutral name "y" to the variable "α". Thus we have

$$\mathrm{Trace}(\mathrm{Frob}_E \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}))$$
$$= \sum_{x,y \text{ in } E} \chi_{2,E}(x(x-1)y(y-1)(x-y)).$$

(8.7.4) Now consider the affine surface S in $\mathbb{A}^3$ over $\mathbb{Z}$ with coordinates x, y, z

$$S : z^2 = x(x-1)y(y-1)(x-y).$$

In order to highlight its symmetry, let us denote by P(t) in $\mathbb{Z}[t]$ the one–variable polynomial

$$P(t) := t(t-1).$$

In terms of P, the equation of S is

$$S : z^2 = P(x)P(y)(x-y).$$

For any finite field E of any odd characteristic p, we have the usual character sum calculation

$$\#S(E) = \sum_{x,y \text{ in } E} \#\{\text{square roots in E of } P(x)P(y)(x-y)\}$$
$$= \sum_{x,y \text{ in } E} (1 + \chi_{2,E}(P(x)P(y)(x-y)))$$
$$= (\#E)^2 + \sum_{x,y \text{ in } E} \chi_{2,E}(P(x)P(y)(x-y))$$
$$= (\#E)^2 + \mathrm{Trace}(\mathrm{Frob}_E \mid H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F})).$$

(8.7.5) Now the sheaf $\mathcal{F}_\ell$ on $\mathbb{P}^1$ makes uniform sense over $\mathbb{Z}[1/2\ell]$: it is lisse on $\mathbb{P}^1 - \{0, 1, \infty\}$, (necessarily) tame along 0, 1, and ∞, and extended by zero to all of $\mathbb{P}^1$. Therefore (cf. [Ka–SE, 4.7.1]), the cohomology groups $H^1(\mathbb{P}^1\otimes\overline{\mathbb{F}}_p, \mathcal{F}_\ell)$ for variable $p \neq 2$ or ℓ are the stalks at the (geometric points over the) closed points of a lisse sheaf $\mathcal{H}_\ell$ on $\mathbb{Z}[1/2\ell]$, whose geometric generic fibre is $H^1(\mathbb{P}^1 \otimes \overline{\mathbb{Q}}, \mathcal{F}_\ell)$. Or in more down to earth language, $H^1(\mathbb{P}^1 \otimes \overline{\mathbb{Q}}, \mathcal{F}_\ell)$ is a two–dimensional $\mathbb{Q}_\ell$–representation $\rho_{\mathrm{gal},\ell}$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ which is unramified outside of 2ℓ, and in which the Frobenius conjugacy classes Frob_p at primes not 2 or ℓ have characteristic polynomials given by

$$\det(1 - T\mathrm{Frob}_p \mid \mathcal{H}_\ell) = \det(1 - T\rho_{\mathrm{gal},\ell}(\mathrm{Frob}_p))$$
$$= \det(1 - T\mathrm{Frob}_E \mid H^1(\mathbb{P}^1 \otimes \overline{\mathbb{F}}_p, \mathcal{F}_\ell)).$$

Moreover, $H^1(\mathbb{P}^1 \otimes \overline{\mathbb{Q}}, \mathcal{F}_\ell)(1)$ is orthogonally self–dual, and pure of weight zero.

(8.7.6) The trace formula above thus says

$$\mathrm{Trace}(\rho_{\mathrm{gal},\ell}(\mathrm{Frob}_p)) = \sum_{x,y \text{ in } E} \chi_{2,E}(P(x)P(y)(x-y)).$$

The right hand side is visibly an integer, independent of ℓ. So the representations $\rho_{\mathrm{gal},\ell}$ form a compatible system of two–dimensional ℓ–adic representations.

(8.7.7) Let us next observe that if $p \equiv 3 \bmod 4$, or more generally if we work over a finite field E in which -1 is not a square, then

$$\sum_{x,y \text{ in } E} \chi_{2,E}(P(x)P(y)(x-y)) = 0.$$

Indeed, interchanging x and y does not change the sum, but changes the sign of $P(x)P(y)(x-y)$. As -1 is a nonsquare in E, this interchange also changes the sign of each term $\chi_{2,E}(P(x)P(y)(x-y))$. Thus the sum is an integer which is equal to minus itself.

(8.7.8) So we have

$$\mathrm{Trace}(\rho_{\mathrm{gal},\ell}(\mathrm{Frob}_p)) = 0 \text{ if } p \equiv 3 \bmod 4,\ p \neq 2 \text{ or } \ell.$$

(8.7.9) Let us view $\rho_{\mathrm{gal},\ell}$ as a two–dimensional $\mathbb{Q}_\ell$–representation of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))$ on $H^1(\mathbb{P}^1 \otimes \overline{\mathbb{Q}}, \mathcal{F}_\ell)$. By Chebotarev, the vanishing of $\mathrm{Trace}(\rho_{\mathrm{gal},\ell}(\mathrm{Frob}_p))$ for $p \equiv 3 \bmod 4$ implies its vanishing outside the entire "Gaussian" subgroup $\pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$ of index two. [Indeed, the function $\gamma \mapsto f(\gamma)$ on $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))$ which is defined as

$$f(\gamma) := \mathrm{Trace}(\rho_{\mathrm{gal},\ell}(\gamma)), \text{ if } \gamma \text{ is in } \pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell])),$$
$$f(\gamma) := 0, \text{ if } \gamma \text{ is not in } \pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell])),$$

is a continuous central function, which agrees with the continuous central function $\gamma \mapsto \mathrm{Trace}(\rho_{\mathrm{gal},\ell}(\gamma))$ on all Frobenii, so these two functions must coincide.]

The Tate–twisted $\mathcal{H}_\ell(1)$, i.e., the representation $\rho_{\mathrm{gal},\ell}(1)$ on $H^1(\mathbb{P}^1 \otimes \overline{\mathbb{Q}}, \mathcal{F})(1)$, is pure of weight zero, and orthogonally self–dual.

Lemma 8.7.10 The representation $\rho_{\mathrm{gal},\ell}(1)$ is irreducible.

proof We first show that $\rho_{\mathrm{gal},\ell}(1)$ is completely reducible. Indeed, consider the Zariski closure G in the orthogonal group O(2) of the image of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))$ under $\rho_{\mathrm{gal},\ell}(1)$. The only

Zariski−closed subgroups of O(2) are O(2), SO(2), and finite groups, all of which are reductive, so the group G is reductive, and hence $\rho_{\mathrm{gal},\ell}(1)$ is completely reducible.

Thus if $\rho_{\mathrm{gal},\ell}(1)$ is reducible, it is (after extension of scalars from $\mathbb{Q}_\ell$ to $\overline{\mathbb{Q}}_\ell$) the direct sum of two characters, say $\sigma \oplus \tau$. For every $p \equiv 3 \bmod 4$, $p \neq \ell$, we saw in 8.7.8 that

$$\mathrm{Trace}(\rho_{\mathrm{gal},\ell}(1)(\mathrm{Frob}_p)) = 0.$$

Thus for every $p \equiv 3 \bmod 4$, $p \neq \ell$, we have

$$\sigma(\mathrm{Frob}_p) = -\tau(\mathrm{Frob}_p).$$

Let us denote by χ_4 the ± 1−valued character of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2]))$ defined by the quadratic extension $\mathbb{Q}(i)/\mathbb{Q}$: concretely, for odd primes p we have

$$\chi_4(\mathrm{Frob}_p) = 1 \text{ if } p \equiv 1 \bmod 4,$$
$$= -1 \text{ if } p \equiv 3 \bmod 4.$$

We observe that $\tau/\sigma = \chi_4$ on $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2]))$. Indeed, $\tau\chi_4/\sigma$ is trivial **outside** the Gaussian subgroup $\pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$ of index two. Therefore $\tau\chi_4/\sigma$ is trivial on all of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))$. [If $H \subset G$ is any proper subgroup of any group, every element h of H is of the form $A^{-1}B$ for two elements A and B in G but not in H: pick any single g not in H, and write $h = g^{-1}(gh)$. So if a linear character of G is trivial outside of H, it is trivial.]

Thus if $\rho_{\mathrm{gal},\ell}(1)$ is reducible, it is of the form $\sigma \oplus \sigma\chi_4$. Because $\rho_{\mathrm{gal},\ell}(1)$ is orthogonal, its determinant has order dividing two. So $\sigma^2\chi_4$ has order dividing two, hence σ^2 has order dividing two, hence σ and $\sigma\chi_4$ each take values which are fourth roots of unity. Therefore, every value of $\mathrm{Trace}(\rho_{\mathrm{gal},\ell}(1))$ lies in $\mathbb{Z}[i]$, and in particular is an algebraic integer. But this is not the case. If $\ell \neq 5$, we readily calculate

$$\mathrm{Trace}(\rho_{\mathrm{gal},\ell}(1)(\mathrm{Frob}_5))$$
$$= (1/5)\textstyle\sum_{\alpha \neq 0,1 \text{ in } \mathbb{F}_5} \sum_{x \text{ in } \mathbb{F}_5} \chi_{2,E}(\alpha(\alpha-1)x(x-1)(x-\alpha))$$
$$= -6/5.$$

If $\ell = 5$, we compute

$$\mathrm{Trace}(\rho_{\mathrm{gal},\ell}(1)(\mathrm{Frob}_{13}))$$
$$= (1/13)\textstyle\sum_{\alpha \neq 0,1 \text{ in } \mathbb{F}_{13}} \sum_{x \text{ in } \mathbb{F}_{13}} \chi_{2,E}(\alpha(\alpha-1)x(x-1)(x-\alpha))$$
$$= 10/13.$$

Therefore $\rho_{\mathrm{gal},\ell}(1)$ is irreducible. QED for 8.7.10

(8.7.11) So $\rho_{\mathrm{gal},\ell}(1)$ is an irreducible orthogonal representation of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))$ of dimension two, whose trace function vanishes outside on the Gaussian subgroup $\pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$. By Theorem 3.5.2, there exists a $\overline{\mathbb{Q}}_\ell$−valued character σ of the Gaussian subgroup $\pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$ such that (after extension of scalars from $\mathbb{Q}_\ell$ to $\overline{\mathbb{Q}}_\ell$)

$$\rho_{gal,\ell}(1) = \mathrm{Ind}(\sigma).$$

The character σ is pure of weight zero. We claim that $\rho_{gal,\ell}(1) | \pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$ is $\sigma \oplus \bar{\sigma}$, for $\bar{\sigma}$ the inverse character to σ. [The notation $\bar{\sigma}$ is slightly abusive: it is only on Frobenii that σ and $\bar{\sigma}$ need take complex conjugate values after any embedding of $\bar{\mathbb{Q}}_\ell$ into $\mathbb{C}$.] To see this, recall from the proof of 3.5.2 that

$$\rho_{gal,\ell}(1) | \pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell])) = \sigma + \tau,$$

for two distinct characters σ and τ of $\pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$. We know that $\sigma + \tau = \bar{\sigma} + \bar{\tau}$ (because $\mathrm{Trace}(\rho_{gal,\ell}(1))$ takes rational values on Frobenii). So either $\tau = \bar{\sigma}$ as asserted, or both σ and τ have order dividing two. In this latter case, $\mathrm{Trace}(\rho_{gal,\ell}(1)) | \pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$ would take only the values 0 and ±2. But we have seen above that the traces of Frob_p for p=5 and p=13 both fail to be algebraic integers. Therefore $\rho_{gal,\ell}(1) | \pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$ is $\sigma \oplus \bar{\sigma}$.

(8.7.12) In particular, $\det(\rho_{gal,\ell}(1))$ is trivial on the Gaussian subgroup $\pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell]))$. On the other hand, $\det(\rho_{gal,\ell}(1))$ is nontrivial: otherwise $\rho_{gal,\ell}(1)$ would have image in the abelian group SO(2), and so would be reducible.

(8.7.13) Therefore $\det(\rho_{gal,\ell}(1))$ is the unique nontrivial character of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))/\pi_1(\mathrm{Spec}(\mathbb{Z}[i,1/2\ell])) = \mathrm{Gal}(\mathbb{Q}(i)/\mathbb{Q})$, i.e., it is the quadratic character of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ cut out by $\mathbb{Q}(i)$. Explicitly, for odd primes $p \neq \ell$, we have

$$\det(\rho_{gal,\ell}(1)(\mathrm{Frob}_p)) = 1 \text{ if } p \equiv 1 \bmod 4,$$
$$= -1 \text{ if } p \equiv 3 \bmod 4. \quad \text{QED}$$

8.8 Explicit determination of the representation $\rho_{gal,\ell}$

(8.8.1) We now explain the numerical coincidence we found in 8.7.4, that for S the affine surface over $\mathbb{Z}$ with equation

$$S : z^2 = x(x-1)y(y-1)(x-y),$$

we had, for every finite field E of odd characteristic, and every prime number ℓ invertible in E, the identity of traces

$$\mathrm{Trace}(\mathrm{Frob}_E | H^1(\mathbb{P}^1\otimes\bar{E}, \mathcal{F}_\ell)) = \#S(E) - (\#E)^2.$$

It has a simple cohomological explanation: it is just the Lefschetz Trace Formula for the surface S.

Lemma 8.8.1.1 Let k be a field in which 2 is invertible, $\bar{k}$ an algebraic closure of k, ℓ a prime invertible in k. The compact cohomology groups $H_c^i(S\otimes\bar{k}, \mathbb{Q}_\ell)$ as $\mathrm{Gal}(\bar{k}/k)$–modules are given by

$$H_c^4(S\otimes\bar{k}, \mathbb{Q}_\ell) \cong \mathbb{Q}_\ell(-2),$$
$$H_c^2(S\otimes\bar{k}, \mathbb{Q}_\ell) \cong H^1(\mathbb{P}^1\otimes\bar{k}, \mathcal{F}_\ell),$$
$$H_c^i(S\otimes\bar{k}, \mathbb{Q}_\ell) = 0 \text{ for all other i.}$$

proof To clarify what is going on, in the equation for S, rename the variables x, y, z as λ, x, y, so S is now the affine surface

$$y^2 = \lambda(\lambda-1)x(x-1)(x-\lambda).$$

View S as sitting over the affine λ line, say

$$f: S \to \mathbb{A}^1,$$
$$(\lambda, x, y) \mapsto \lambda.$$

Consider the Leray spectral sequence

$$E_2^{a,b} = H_c^a(\mathbb{A}^1\otimes\overline{k}, R^bf_!\mathbb{Q}_\ell) \Rightarrow H_c^{a+b}(S\otimes\overline{k}, \mathbb{Q}_\ell).$$

Over the open set $\mathbb{A}^1[1/\lambda(\lambda-1)]$ of the base, the induced map

$$f : S[1/\lambda(\lambda-1)] \to \mathbb{A}^1[1/\lambda(\lambda-1)]$$

is $\mathcal{E} - \{0\} \to | \mathbb{A}^1[1/\lambda(\lambda-1)]$, for $\pi : \mathcal{E} \to \mathbb{A}^1[1/\lambda(\lambda-1)]$ the twisted Legendre family. Since removing a single point from a projective smooth geometrically connected curve does not change its H_c^1 or its H_c^2, we have

$$R^1f_!\mathbb{Q}_\ell | \mathbb{A}^1[1/\lambda(\lambda-1)] = R^1\pi_!\mathbb{Q}_\ell | \mathbb{A}^1[1/\lambda(\lambda-1)]$$
$$= R^1\pi_*\mathbb{Q}_\ell | \mathbb{A}^1[1/\lambda(\lambda-1)] = \mathcal{F}_\ell | \mathbb{A}^1[1/\lambda(\lambda-1)],$$

and

$$R^2f_!\mathbb{Q}_\ell | \mathbb{A}^1[1/\lambda(\lambda-1)] \cong \mathbb{Q}_\ell(-1).$$

Since an affine smooth curve has vanishing H_c^0, proper base change gives us

$$R^0f_!\mathbb{Q}_\ell | \mathbb{A}^1[1/\lambda(\lambda-1)] = 0.$$

Over the points λ=0 and λ=1, the fibre of f is the (nonreduced) affine curve in x,y space with equation $y^2 = 0$. But etale cohomology does not see nilpotents, so these special fibres might as well be $\mathbb{A}^1$'s, whose H_c^0 and H_c^1 both vanish, and H_c^2 is $\mathbb{Q}_\ell(-1)$.

Denote by $j : \mathbb{A}^1[1/\lambda(\lambda-1)] \to \mathbb{A}^1$ the inclusion. Proper base change gives

$$R^0f_!\mathbb{Q}_\ell = 0 \text{ on } \mathbb{A}^1,$$
$$R^1f_!\mathbb{Q}_\ell = j_!j^*R^1f_!\mathbb{Q}_\ell = j_!j^*\mathcal{F}_\ell.$$

The sheaf $\mathcal{F}_\ell$ also vanishes over λ=0 and λ=1, so we have

$$R^1f_!\mathbb{Q}_\ell = \mathcal{F}_\ell | \mathbb{A}^1.$$

As the sheaf $\mathcal{F}_\ell$ also vanishes over the point $\lambda=\infty$ in $\mathbb{P}^1$, we have

$$H^1(\mathbb{P}^1\otimes\overline{k}, \mathcal{F}_\ell) = H_c^1(\mathbb{A}^1\otimes\overline{k}, \mathcal{F}_\ell | \mathbb{A}^1).$$

Thus we have

$$H^1(\mathbb{P}^1\otimes\overline{k}, \mathcal{F}_\ell) = H_c^1(\mathbb{A}^1\otimes\overline{k}, R^1f_!\mathbb{Q}_\ell).$$

All the geometric fibres of f, when reduced, are irreducible curves, so we have

$$R^2 f_! \mathbb{Q}_\ell \cong \mathbb{Q}_\ell(-1).$$

With this data in hand, we easily compute the E_2 terms in the spectral sequence. All the sheaves $R^i f_! \mathbb{Q}_\ell$ on $\mathbb{A}^1$ are middle extensions on an affine smooth curve, so we have

$$E_2^{0,b} = 0 \text{ for all b.}$$

Among all the sheaves $R^i f_! \mathbb{Q}_\ell$, only $R^1 f_! \mathbb{Q}_\ell$ ($\cong \mathcal{F}_\ell \mid \mathbb{A}^1$) is not geometrically constant. As $H_c^1(\mathbb{A}^1 \otimes \overline{k}, \mathbb{Q}_\ell)$ vanishes, we have

$$E_2^{1,b} = 0 \text{ for } b \neq 1,$$
$$E_2^{1,1} = H^1(\mathbb{P}^1 \otimes \overline{k}, \mathcal{F}_\ell).$$

The sheaf $R^1 f_! \mathbb{Q}_\ell$ ($\cong \mathcal{F}_\ell \mid \mathbb{A}^1$) is an irreducible middle extension of rank 2, so its H_c^2 vanishes, and so we find

$$E_2^{2,b} = 0 \text{ for } b \neq 2,$$
$$E_2^{2,2} = H_c^2(\mathbb{A}^1 \otimes \overline{k}, \mathbb{Q}_\ell(-1)) \cong \mathbb{Q}_\ell(-2).$$

With such a paucity of nonzero E_2 terms, the spectral sequence degenerates at E_2, and gives the asserted values for the compact cohomology groups of S. QED

(8.8.2) When we view $H^1(\mathbb{P}^1 \otimes \overline{k}, \mathcal{F}_\ell)$ as $H_c^2(S \otimes \overline{k}, \mathbb{Q}_\ell)$, the cup–product pairing

$$H^1(\mathbb{P}^1 \otimes \overline{k}, \mathcal{F}_\ell) \times H^1(\mathbb{P}^1 \otimes \overline{k}, \mathcal{F}_\ell) \to \mathbb{Q}_\ell(-2)$$

becomes the cup–product pairing

$$H_c^2(S \otimes \overline{k}, \mathbb{Q}_\ell) \times H_c^2(S \otimes \overline{k}, \mathbb{Q}_\ell) \to H_c^4(S \otimes \overline{k}, \mathbb{Q}_\ell) \cong \mathbb{Q}_\ell(-2).$$

Since the pairing on $H^1(\mathbb{P}^1 \otimes \overline{k}, \mathcal{F}_\ell)$ is nondegenerate, we find that the cup–product pairing on $H_c^2(S \otimes \overline{k}, \mathbb{Q}_\ell)$ is nondegenerate. Since S is an affine and singular surface, this nondegeneracy seems highly non–obvious.

(8.8.3) As we saw in the proof of 8.8.1.1, $S[1/\lambda(\lambda-1)]$ is $\mathcal{E} - \{0\}$ for

$$\pi : \mathcal{E} \to \mathbb{A}^1[1/\lambda(\lambda-1)]$$

the twisted Legendre family, whose affine equation is

$$y^2 = \lambda(\lambda - 1)x(x - 1)(x - \lambda).$$

(8.8.4) There is a canonical way to complete $\pi : \mathcal{E} \to \mathbb{A}^1[1/\lambda(\lambda-1)]$ to an elliptic surface

$$\overline{\pi} : \mathbb{E} \to \mathbb{P}^1,$$

(i.e., $\mathbb{E}$ is a projective smooth geometrically connected surface, and $\overline{\pi}$ coincides with π over $\mathbb{P}^1 - \{0, 1, \infty\}$) in such a way that the fibres over the three points 0, 1, ∞ are the Kodaira–Néron special fibres of the elliptic curve $y^2 = \lambda(\lambda - 1)x(x - 1)(x - \lambda)$ considered successively over the complete fields $k((\lambda))$, $k((1-\lambda))$, and $k((1/\lambda))$.

(8.8.5) Over each of these fields, this curve is of type $\mathrm{I}^*{}_2$. [Over $k((\lambda))$ we rewrite the equation as

$$(\lambda y)^2 = (\lambda-1)(\lambda x)(\lambda x - \lambda)(\lambda x - \lambda^2),$$

so in new variables $X = -\lambda x$ and $Y = \lambda y/\mathrm{Sqrt}(1-\lambda)$ we have

$$Y^2 = X(X + \lambda)(X + \lambda^2).$$

Over $k((t))$ with t either $1-\lambda$ or $1/\lambda$, similar changes of variable bring our curve to the form

$$Y^2 = X(X - t)(X - t^2).]$$

The Tate algorithm [Sil–ATEC, page 366] shows that over each of 0, 1, ∞, the special fibre consists of seven $\mathbb{P}^1$'s over k, of which four are reduced and three have multiplicity two, with a total of six crossing points, arranged as

```
        *                        *
*******************************
        *                        *
***********          ************
        *                        *
***********          ************
        *                        *
```

(8.8.6) Suppose we start over $\mathbb{F}_p$, for an odd prime p, and pick a prime $\ell \neq p$. Then over any finite field k of characteristic p, we have

(8.8.6.1)
$$\begin{aligned}
\#\mathbb{E}(k) &= \#((\pi^{-1}\{0, 1, \infty\})(k)) + \#(\pi^{-1}(\mathbb{A}^1[1/\lambda(\lambda-1)])(k)) \\
&= 3(7(\#k) + 1) + \#\mathcal{E}(k) \\
&= 3(7(\#k) + 1) + \#(\mathbb{A}^1[1/\lambda(\lambda-1)](k)) + \#(\mathcal{E} - \{0\})(k) \\
&= 3(7(\#k) + 1) + \#(\mathbb{A}^1[1/\lambda(\lambda-1)](k)) + \#(S[1/\lambda(\lambda-1)](k)) \\
&= 3(7(\#k) + 1) + (\#k - 2) + \#S(k) - 2(\#k) \\
&=20\#k + 1 + \#S(k) \\
&=20\#k + 1 + (\#k)^2 + \mathrm{Trace}(\mathrm{Frob}_k \mid H^1(\mathbb{P}^1\otimes\bar{k}, \mathcal{F}_\ell)).
\end{aligned}$$

(8.8.7) Using the Weil Conjectures, we infer that the Betti numbers of $\mathbb{E}$ are 1, 0, 22, 0, 1.

(8.8.8) On the other hand, the minimal projective nonsingular model of the affine surface S is a K3 surface. Indeed, it is the K3 surface "X_4", which is the (minimal resolution of the) double covering of $\mathbb{P}^2$ branched along $XYZ(X-Y)(X-Z)(Y-Z)$, cf. [Beu–St, page 283, case $\mathcal{A}$]. Being a K3 surface, X_4 is an absolutely minimal model of its function field. What is the relation between $\mathbb{E}$ and X_4? Since $\mathbb{E}$ is also a projective nonsingular model of S, the tautological birational map from $\mathbb{E}$ to X_4 is, by the absolute minimality of X_4, a morphism. Any birational morphism between projective smooth surfaces is a succession of blowings up of points. But $\mathbb{E}$ has middle Betti number 22, the same as the K3 surface X_4, so there can be no blowings up. Thus $\mathbb{E} \cong X_4$.

(8.8.9) According to Beukers and Stienstra [Beu–St, page 292], elaborating a theorem of Shioda and Inose [Shio–In, Thm. 6], for any odd prime p the zeta function of $X_4/\mathbb{F}_p$ is equal to

$$1/(1 - T)P_p(T)(1 - pT)^{20}(1-p^2T)$$

for $P_p(T)$ the quadratic polynomial given by

$1 - 2(a^2 - b^2)T + p^2T^2$, if $p \equiv 1 \bmod 4$, $p = a^2 + b^2$, a odd,

$1 - p^2T^2$, if $p \equiv 3 \bmod 4$.

In particular, $\#X_4(\mathbb{F}_p)$ is given by:

$1 + 20p + p^2 + 2(a^2 - b^2)$, if $p \equiv 1 \bmod 4$, $p = a^2 + b^2$, a odd,

$1 + 20p + p^2$, if $p \equiv 3 \bmod 4$.

Comparing with our formulas for $\#\mathbb{E}(\mathbb{F}_p)$ in 8.8.6.1, we find

(8.8.9.1) $\quad \mathrm{Trace}(\mathrm{Frob}_k \mid H^1(\mathbb{P}^1\otimes\bar{k}, \mathcal{F}_\ell)) = 2(a^2 - b^2)$, if $p = a^2 + b^2$, with a odd,

$\mathrm{Trace}(\mathrm{Frob}_k \mid H^1(\mathbb{P}^1\otimes\bar{k}, \mathcal{F}_\ell)) = 0$, if $p \equiv 3 \bmod 4$.

(8.8.10) These explicit formulas have a simple meaning in terms of the representation ρ_{gal}. Denote by ρ_4 the grossencharacter of $\mathbb{Q}(i)$ of conductor 2+2i attached to the elliptic curve $y^2 = x^3 - x$, given explicitly on ideals of $\mathbb{Z}[i]$ which are prime to 2 by the formula $\chi_4(I) = \alpha$ where α is the unique generator of the ideal I which satisfies $\alpha \equiv 1 \bmod 2+2i$. Fix a prime ℓ and an embedding of $\mathbb{Q}(i)$ into $\bar{\mathbb{Q}}_\ell$. Then ρ_4 gives rise to a $\bar{\mathbb{Q}}_\ell$–valued character $\rho_{4,\ell}$ of $\pi_1(\mathrm{Spec}(\mathbb{Z}[i, 1/2\ell])$ with the following property. For each Gaussian prime π not dividing 2ℓ, with $\pi \equiv 1 \bmod 2+2i$, we have

$$\rho_{4,\ell}(\mathrm{Frob}_\pi) = \pi.$$

Proposition 8.8.11 The two–dimensional representation $\rho_{\mathrm{gal},\ell}$ of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))$ afforded by $H^1(\mathbb{P}^1\otimes\bar{\mathbb{Q}}, \mathcal{F}_\ell)$, is $\mathrm{Ind}((\rho_{4,\ell})^2)$, the induction of $(\rho_{4,\ell})^2$ from $\pi_1(\mathrm{Spec}(\mathbb{Z}[i, 1/2\ell]))$ to $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell]))$.

proof We have shown in 8.7.10 above that $\rho_{\mathrm{gal},\ell}$ is irreducible. Hence ρ_{gal} is semisimple. The induction of a linear character (or of any semisimple $\bar{\mathbb{Q}}_\ell$–representation) from a subgroup of finite index is semisimple. So $\mathrm{Ind}((\rho_{4,\ell})^2)$ is a semisimple representation. The two representations $\rho_{\mathrm{gal},\ell}$ and $\mathrm{Ind}((\rho_{4,\ell})^2)$ of $\pi_1(\mathrm{Spec}(\mathbb{Z}[1/2\ell])$ have the same trace function on all Frobenius elements, by 8.8.9.1. By Chebotarev, their trace functions are equal. Hence these two representations have isomorphic semisimplifications. As both representations are semisimple, they are isomorphic. QED

8.9 A family of interesting examples

(8.9.1) Let us return to the situation of 8.6. Thus k is a finite field of odd characteristic, ℓ is a prime number invertible in k, and over $\mathbb{P}^1 - \{0, 1, \infty\}$ with parameter λ we consider the twisted Legendre family of elliptic curves

$$\pi : \mathcal{E} \to \mathbb{P}^1 - \{0, 1, \infty\},$$

given by the affine equation

$$y^2 = \lambda(\lambda - 1)x(x - 1)(x - \lambda).$$

We denote by $j : \mathbb{P}^1 - \{0, 1, \infty\} \to \mathbb{P}^1$ the inclusion. For each **odd** integer $n \geq 1$, we consider the lisse sheaf

$$\mathcal{F}_1 := R^1\pi_*\overline{\mathbb{Q}}_\ell$$

on $\mathbb{P}^1 - \{0, 1, \infty\}$. Then $\mathcal{F}_1$ is lisse of rank 2, pure of weight one, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-1)$. Along the sections 0, 1, and ∞ of C/T, the local monodromy of $\mathcal{F}$ is

(the quadratic character)⊗(unipotent nontrivial).

For each **odd** integer $m \geq 1$, take $\mathcal{F}_m := \mathrm{Sym}^m(\mathcal{F}_1)$. Thus $\mathcal{F}_m$ is lisse of even rank m+1, pure of weight m, and orthogonally self–dual toward $\overline{\mathbb{Q}}_\ell(-m)$. Because $\mathcal{F}_1$ has $G_{geom} = SL(2)$, $\mathcal{F}_m$ is geometrically irreducible. Its local monodromy along the sections 0, 1, ∞ is

(the quadratic character)⊗(a single unipotent Jordan block).

Thus $\mathcal{F}_m$ is everywhere tame, and at each of its singularities, the dimension of $\mathcal{F}_m/(\mathcal{F}_m)^I$ is the even integer m+1.

(8.9.1.1) For each **even** integer $d \geq 144$, we define a divisor D_d in $\mathbb{P}^1$ by $D_d := d\infty$, and form the lisse sheaf

$$\mathcal{G}_{d,m} := \mathrm{Twist}_{\chi_2,\mathbb{P}^1,D_d}(j_*\mathcal{F}_m)$$

on the space

$$X_d := \mathit{Fct}(\mathbb{P}^1, D, d, \{0,1\})$$

of degree d polynomials in λ with invertible discriminant and which are invertible at both 0 and 1. The Tate–twisted sheaf $\mathcal{G}_{d,m}((m+1)/2)$ is orthogonally self–dual, of rank (m+1)(d+1). According to Theorem 8.5.7, part 1), for d >> 0, the group G_{geom} for $\mathcal{G}_{d,m}((m+1)/2)$ is SO((m+1)(d+1)). Moreover, for any such d, the group G_{arith} for $\mathcal{G}_{d,m}((m+1)/2)$ is SO((m+1)(d+1)) if and only if the sign in the functional equation for the L–function of $j_*\mathcal{F}_m((m+1)/2)$ on $\mathbb{P}^1\otimes k$ is +1. In 8.6.5 we determined this sign for the case m=1 by a global number field argument. Here we give a different proof, based on the theory of local constants, which works for all m.

Theorem 8.9.2 Hypotheses and notations as above, for any finite field k of odd characteristic, any prime number ℓ invertible in k, and any odd integer $m \geq 1$, the sign in the functional equation for the L–function of $j_*\mathcal{F}_m((m+1)/2)$ on $\mathbb{P}^1\otimes k$ is given by

$$\det(-\mathrm{Frob}_k \mid H^1(\mathbb{P}^1\otimes\overline{k}, j_*\mathcal{F}_m((m+1)/2))) = \chi_2(-1)^{(m+1)/2}.$$

[Recall that $\chi_2(-1)$ is equal to

+1, if $\#k \equiv 1 \bmod 4$,
−1, if $\#k \equiv 3 \bmod 4$.]

proof Since we are trying to determine a sign, and no power of #k is a root of unity, we may work in the multiplicative group $(\overline{\mathbb{Q}}_\ell)^\times/(\#k)^{\mathbb{Z}}$. We write $a \approx b$ if a/b is an integer power of #k. By [De–Const, 7.9], valid here because $\mathcal{F}_m$ is part of a compatible system, the constant in the functional equation is given by

$$1/\det(-\mathrm{Frob}_k \mid H^1(\mathbb{P}^1\otimes\overline{k}, j_*\mathcal{F}_m((m+1)/2)))$$
$$= \prod_{v \text{ in } \mathbb{P}^1} \varepsilon(V_{m,v}, \psi_v, \mu_v),$$

the product over the closed points v of $\mathbb{P}^1\otimes k$. Here $V_{m,v}$ denotes the restriction of $\mathcal{F}_m((m+1)/2)$ to the decomposition group D_v at v, and ψ_v and μ_v are the local components of a nontrivial additive character ψ of, and of Haar measure μ of total mass one on, the quotient additive group A_K/K of the adeles A_K of $K := k(\lambda)$ by the discrete subgroup K. We can make these choices so that μ_v gives the integer ring $\mathcal{O}_v$ total mass one for all but finitely many v, and gives it mass an integer power of #k for every v. We get an explicit choice of ψ as follows. Pick a nonzero meromorphic one–form ω on $\mathbb{P}^1\otimes k$, and a nontrivial additive character ψ_0 of k. Then we get a global ψ by defining $\psi_x(f) := \psi_0(\mathrm{Trace}_{k(v)/k}(\mathrm{Res}_v(f\omega)))$. We will choose ω so that it has simple poles at each of 0, 1, ∞, with residue +1 at each.

With these choices, we first claim that for each v other than 0, 1, ∞, we have

$$\varepsilon(V_{m,v}, \psi_v, \mu_v) \approx 1.$$

At such v, $V_{m,v}$ is unramified of even rank m+1, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(1)$. So $\det(V_{m,v}) \cong \overline{\mathbb{Q}}_\ell((m+1)/2)$. By the transformation formulas [De–Const, 5.3 and 5.4], the constant $\varepsilon(V_{m,v}, \psi_v, \mu_v)$ is, up to $\approx$ equivalence, independent of the choice of measure μ_v giving $\mathcal{O}_v$ mass an integer power of #k, and independent of the choice of local character ψ_v. Choose ω to have neither zero nor pole at v, and choose μ_v to give $\mathcal{O}_v$ mass one. Then $\varepsilon(V_{m,v}, \psi_v, \mu_v) = 1$. This follows from [De–Const, 5.5.3], applied with its W taken to be $V_{m,v}$ and its V taken to be $\mathbb{1}$, and [De–Const, 5.9], applied with its χ taken to be $\mathbb{1}$.

At v any of the three points 0, 1, ∞, $V_{m,v}$ has even rank m+1, and is symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(1)$. So $\det(V_{m,v}) \cong \overline{\mathbb{Q}}_\ell((m+1)/2)$. By [De–Const,5.3 and 5.4], $\varepsilon(V_{m,v}, \psi_v, \mu_v)$ is, up to $\approx$ equivalence, independent of the choice μ_v giving $\mathcal{O}_v$ measure an integer power of #k, and independent of the choice of local character ψ_v.

For odd $m \geq 1$, we have

$$V_{m,v} = \mathrm{Sym}^m(V_{1,v})((1-m)/2).$$

So we have

$$\varepsilon(V_{m,v}, \psi_v, \mu_v) \approx \varepsilon(\mathrm{Sym}^m(V_{1,v}), \psi_v, \mu_v).$$

Now $V_{1,v}(-1)$ is just the H^1 of the twisted Legendre curve

$$y^2 = \lambda(\lambda - 1)x(x - 1)(x - \lambda),$$

viewed as a representation of the decomposition group D_v. Over $k((\lambda))$, $1-\lambda$ is a square, and the twisted Legendre curve is isomorphic to

$$y^2 = (-\lambda)x(x - 1)(x - \lambda).$$

Over $k((1-\lambda))$, $\lambda = 1 - (1-\lambda)$ is a square, and the twisted Legendre curve is isomorphic to

$$y^2 = (\lambda-1)x(x - 1)(x - \lambda).$$

Over $k((1/\lambda))$, $\lambda(\lambda-1)$ is a square, and the twisted Legendre curve is isomorphic to

$$y^2 = x(x - 1)(x - \lambda).$$

Now consider the Legendre curve itself,

$$y^2 = x(x - 1)(x - \lambda),$$

over $k(\lambda)$. One sees from the Tate algorithm [Sil–ATEC, page 366] that it has split multiplicative reduction of type I_2 at $\lambda=1$, so its H^1 has unipotent local monodromy at $\lambda=1$, and as a representation of D_1 it has $(H^1)^I \cong \overline{\mathbb{Q}}_\ell$, and $H^1/(H^1)^I \cong \overline{\mathbb{Q}}_\ell(-1)$. Now $V_{1,1}(-1)$ as representation of D_1 is $\mathcal{L}_{\chi_2(\lambda-1)}\otimes$(this H^1), so $V_{1,1}(-1)$ as D_1–representation is an extension of the two characters

$$\mathcal{L}_{\chi_2(\lambda-1)}, \mathcal{L}_{\chi_2(\lambda-1)}(-1).$$

At $\lambda=0$, one sees from the Tate algorithm [Sil–ATEC, page 366] that the Legendre curve has multiplicative reduction of type I_2, and this reduction is split if and only if -1 is a square in k. So its H^1 has unipotent local monodromy at $\lambda=0$, and as a representation of D_0 it has

$$(H^1)^I \cong \mathcal{L}_{\chi_2(-1)}, \text{ and } H^1/(H^1)^I \cong \mathcal{L}_{\chi_2(-1)}(-1).$$

Now $V_{1,0}(-1)$ as representation of D_0 is $\mathcal{L}_{\chi_2(-\lambda)}\otimes$(this H^1), so $V_{1,0}(-1)$ as D_0–representation is an extension of the two characters

$$\mathcal{L}_{\chi_2(\lambda)}, \mathcal{L}_{\chi_2(\lambda)}(-1).$$

At $\lambda=\infty$, take $t := 1/\lambda$ as uniformizing parameter. In the new x, y variables tx and t^2y, the Legendre curve becomes

$$y^2 = tx(x - 1)(x - t).$$

Thus the Legendre curve over $k((1/\lambda))$ is the $-t = -1/\lambda$ twist of a curve with split multiplicative reduction of type I_2 at $\lambda=\infty$. As already noted, our twisted curve is isomorphic to the Legendre curve over $k((1/\lambda))$. So $V_{1,\infty}(-1)$ as D_∞–representation is an extension of the two characters

$$\mathcal{L}_{\chi_2(-1/\lambda)}, \mathcal{L}_{\chi_2(-1/\lambda)}(-1).$$

So for each odd $m \geq 1$, $V_{m,v}$ is a successive extension of various Tate twists of the single character

$$\mathcal{L}_{\chi_2(\lambda)}, \text{ at } v=0,$$
$$\mathcal{L}_{\chi_2(\lambda-1)}, \text{ at } v=1,$$
$$\mathcal{L}_{\chi_2(-1/\lambda)}, \text{ at } v=\infty.$$

The key point is that each of these characters is **ramified**. So at v any of 0, 1, ∞, our local ε constants are equal to the local ε_0 constants. Local ε_0 constants (but not in general the local ε constants) are multiplicative in short exact sequences, cf. [Lau–TFC, 3.1.5.7]. So in the notations of [De–Const, 8.12] or [Lau–TFC, 3.1.5.6–7], we have

$$\varepsilon(V_{m,v}, \psi_v, \mu_v) = \varepsilon_0(V_{m,v}, \psi_v, \mu_v)$$
$$\approx \varepsilon_0(\mathcal{L}_{\chi_2(\lambda)}, \psi_v, \mu_v)^{m+1} \approx \varepsilon(\mathcal{L}_{\chi_2(\lambda)}, \psi_v, \mu_v)^{m+1} \text{ at } v=0.$$

Similarly, we have

$$\varepsilon(V_{m,v}, \psi_v, \mu_v) \approx \varepsilon(\mathcal{L}_{\chi_2(\lambda-1)}, \psi_v, \mu_v)^{m+1}, \text{ at } v=1,$$
$$\varepsilon(V_{m,v}, \psi_v, \mu_v) \approx \varepsilon(\mathcal{L}_{\chi_2(-1/\lambda)}, \psi_v, \mu_v)^{m+1}, \text{ at } v=\infty.$$

Denote by $G(\chi_2, \psi_0)$ the quadratic Gauss sum for k:

$$G(\chi_2, \psi_0) := \Sigma_{x \text{ in } k^\times} \chi_2(x)\psi_0(x).$$

For ψ_v given by an ω with a simple pole at v with residue 1, and μ_v giving $\mathcal{O}_v$ mass #k, we have [De–Const, 5.10.1–2]

$$\varepsilon(\mathcal{L}_{\chi_2(-\lambda)}, \psi_v, \mu_v) = -G(\chi_2, \psi_0) \text{ at } v=0,$$
$$\varepsilon(\mathcal{L}_{\chi_2(1-\lambda)}, \psi_v, \mu_v) = - G(\chi_2, \psi_0) \text{ at } v=1,$$
$$\varepsilon(\mathcal{L}_{\chi_2(1/\lambda)}, \psi_v, \mu_v) = - G(\chi_2, \psi_0) \text{ at } v=\infty.$$

Thus we find

$$1/\det(-\mathrm{Frob}_k \mid H^1(\mathbb{P}^1\otimes\bar{k}, j_*\mathcal{F}_m((m+1)/2)))$$
$$\approx (- G(\chi_2, \psi_0))^{3(m+1)}.$$
$$= (G(\chi_2, \psi_0)^2)^{3(m+1)/2}$$
$$= (\chi_2(-1)(\#k))^{3(m+1)/2}$$
$$\approx \chi_2(-1)^{3(m+1)/2}$$
$$\approx \chi_2(-1)^{(m+1)/2}. \qquad \text{QED}$$

8.10 Another family of examples

(8.10.1) In this section, we work over a finite field k in which 6 is invertible. Fix δ in $k^\times$, and denote by $\mathcal{M}_{\delta,k}$ the affine curve over k in (g_2, g_3)–space defined by the equation

$$\mathcal{M}_{\delta,k} : (g_2)^3 - 27(g_3)^2 = \delta.$$

Over $\mathcal{M}_{\delta,k}$, we have the family of elliptic curves

$$\pi : \mathcal{E} \to \mathcal{M}_{\delta,k},$$

with $\mathcal{E} - \{0\}$ given by the affine equation

$$y^2 = 4x^3 - g_2x - g_3.$$

The pair $(\mathcal{E}, \omega := dx/2y)$ over $\mathcal{M}_{\delta,k}$ is the universal elliptic curve with differential (E, ω) over a k–scheme with $\Delta(E, \omega) = \delta$, cf. [Ka–Maz, 10.13.3].

(8.10.2) The moduli space $\mathcal{M}_{\delta,k}$ is itself the complement of the origin in an elliptic curve $E_{\delta,k}$. We denote by

$$j : \mathcal{M}_{\delta,k} = E_{\delta,k} - \{0\} \to E_{\delta,k}$$

the inclusion. In the following discussion, we will often refer to the origin of $E_{\delta,k}$ as the point at ∞ of $\mathcal{M}_{\delta,k}$.

(8.10.3) Fix a prime number ℓ invertible in k, and form the lisse rank two $\overline{\mathbb{Q}}_\ell$–sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell$ on $\mathcal{M}_{\delta,k}$. This sheaf has as its G_{geom} the group SL(2), because the curve $\mathcal{E}/\mathcal{M}_{\delta,k}$ has nonconstant j–invariant (namely $j = 1728(g_2)^3/\delta$) which has a pole of order six at ∞. The reduction type at ∞ is easily checked to be $I^*{}_6$. After we quadratically twist this curve by $-g_2/2g_3$, it is of type I_6, with split multiplicative reduction at ∞. So $R^1\pi_*\overline{\mathbb{Q}}_\ell$ as representation of the inertia group $I(\infty)$ at ∞ (remember, ∞ is the origin on $E_{\delta,k}$) is

$$\mathcal{L}_{\chi_2(-g_2/2g_3)}\otimes\mathrm{Unip}(2).$$

As a representation of the decomposition group $D(\infty)$ at ∞, $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is an extension of the two characters

$$\mathcal{L}_{\chi_2(-g_2/2g_3)}, \mathcal{L}_{\chi_2(-g_2/2g_3)}(-1).$$

(8.10.4) For any odd integer $m \geq 1$, the sheaf $\mathrm{Sym}^m(R^1\pi_*\overline{\mathbb{Q}}_\ell)$ on $\mathcal{M}_{\delta,k}$ is lisse of rank m+1, pure of weight m, geometrically irreducible, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-m)$. As representation of $I(\infty)$, it is

$$\mathcal{L}_{\chi_2(-g_2/2g_3)}\otimes\mathrm{Unip}(m+1).$$

As a representation of the decomposition group $D(\infty)$ at ∞, it is an extension of the m+1 characters

$$\mathcal{L}_{\chi_2(-g_2/2g_3)}, \mathcal{L}_{\chi_2(-g_2/2g_3)}(-1), ..., \mathcal{L}_{\chi_2(-g_2/2g_3)}(-m).$$

Theorem 8.10.5 Hypotheses and notations as above, for any finite field k of odd characteristic, any prime number ℓ invertible in k, and any odd integer $m \geq 1$, the sign in the functional equation of the L–function of $j_*\mathrm{Sym}^m(R^1\pi_*\overline{\mathbb{Q}}_\ell)((m+1)/2)$ on $E_{\delta,k}$ is given by

$$\det(-\mathrm{Frob}_k \mid H^1(\mathcal{M}_{\delta,k}\otimes\bar{k}, j_*\mathrm{Sym}^m(R^1\pi_*\bar{\mathbb{Q}}_\ell)((m+1)/2)))$$
$$= \chi_2(-1)^{(m+1)/2},$$
$$= +1, \text{ if } \#k \equiv 1 \bmod 4,$$
$$= (-1)^{(m+1)/2}, \text{ if } \#k \equiv 3 \bmod 4.$$

proof The proof is entirely similar to the proof of Theorem 8.9.2. QED

Corollary 8.10.6 Fix a strictly increasing sequence of positive even integers $0 < d_1 < d_2$ .... For each ν, denote by D_ν the divisor $d_\nu\infty$ on $E_{\delta,k}$ (remember, ∞ is the origin on $E_{\delta,k}$). Fix an odd integer $m \geq 1$. Form the twist sheaf

$$\mathcal{G}_{\nu,m} := \mathrm{Twist}_{\chi_2, E_{\delta,k}, D_\nu}(j_*\mathrm{Sym}^m(R^1\pi_*\bar{\mathbb{Q}}_\ell)((m+1)/2))$$

on the space $X_\nu := \mathit{Fct}(E_{\delta,k}, D_\nu, \emptyset)$. This sheaf is lisse of rank $(m+1)(d_\nu + 1)$, pure of weight zero, and orthogonally self–dual. For each $d_\nu \geq 72(m+1)$, G_{geom} for $\mathcal{G}_{\nu,m}$ is $SO((m+1)(d_\nu + 1))$, and G_{arith} for $\mathcal{G}_{\nu,m}$ is

$SO((m+1)(d_\nu + 1))$, if -1 is a square in k,
$O((m+1)(d_\nu + 1))$, if -1 is **not** a square in k.

proof For each odd m, $j_*\mathrm{Sym}^m(R^1\pi_*\bar{\mathbb{Q}}_\ell)((m+1)/2)$ is lisse of even rank m+1 outside of the point ∞, where it is tame, and its inertial invariants vanish. The assertion then follows from Theorem 8.5.7, part 1), applied to $j_*\mathrm{Sym}^m(R^1\pi_*\bar{\mathbb{Q}}_\ell)((m+1)/2)$, and the preceding theorem, which gives the sign in its functional equation. QED

Chapter 9: Twisting by "Primes", and Working over $\mathbb{Z}$

9.0 Construction of some S_d torsors

(9.0.1) In this section, we work over an arbitrary scheme T, which will play the role of a parameter scheme in what follows. We fix a proper, smooth, geometrically connected curve C/T of genus g, and an integer $d \geq 2g+1$. We denote by $\mathrm{Jac}^d(C/T)$, or simply Jac^d, the open and closed subscheme of $\mathrm{Pic}_{C/T}$ formed by divisor classes of degree d. We denote by $\mathrm{Div}^d(C/T)$ the space of **effective** divisors in C of degree d. Thus for any T–scheme Y, a Y–valued point of $\mathrm{Div}^d(C/T)$ is a closed subscheme of $C\times_T Y$ which is finite and locally free over Y of rank d. The scheme $\mathrm{Div}^d(C/T)$ is naturally isomorphic to the scheme $\mathrm{Sym}^d(C/T)$, the quotient of C^d, the d–fold fibre product of C with itself over T, by the symmetric group S_d, cf. [SGA 4, Exposé XVII, 6.3.9]. We have natural morphisms

$$C^d \to \mathrm{Div}^d(C/T) \to \mathrm{Jac}^d(C/T)$$

of smooth T–schemes. The first map is finite and flat of rank d!, and the second map is a $\mathbb{P}^{d-g}$ bundle.

(9.0.2) We denote by

$$\mathrm{EtaleDiv}^d(C/T) \subset \mathrm{Div}^d(C/T)$$

the open subscheme of $\mathrm{Div}^d(C/T)$ whose Y–valued points are the closed subschemes of $C\times_T Y$ which are finite etale over Y of rank d. [More concretely, if T is the spec of an algebraically closed field k, the k–valued points of $\mathrm{EtaleDiv}^d(C/T)$ are the effective divisors of degree d which consist of d distinct points.]

(9.0.3) We denote by

$$(C^d)_{\text{all dist}} \subset (C/T)^d$$

the open subscheme of C^d whose Y–valued points are those d–tuples of points Q_i in C(Y) which are pairwise disjoint, i.e., for each $1 \leq i < j \leq d$, the scheme–theoretic intersection $Q_i \cap Q_j$ in $C\times_T Y$ is empty. Thus we have a cartesian diagram

$$\begin{array}{ccc} (C^d)_{\text{all dist}} & \subset & C^d \\ \downarrow & & \downarrow \\ \mathrm{EtaleDiv}^d(C/T) & \subset & \mathrm{Div}^d(C/T). \end{array}$$

The first vertical map above,

$$(C^d)_{\text{all dist}} \to \mathrm{EtaleDiv}^d(C/T)$$

is a finite etale S_d–torsor.

(9.0.4) Now suppose we are given an effective relative Cartier divisor Z in C. We denote by

$$\mathrm{EtaleDiv}^d(C/T, Z) \subset \mathrm{EtaleDiv}^d(C/T)$$

the open subscheme of $\mathrm{EtaleDiv}^d(C/T)$ whose Y−valued points are the closed subschemes of $C\times_T Y$ which are finite etale over T of rank d and disjoint from $Z\times_T Y$. [More concretely, the geometric points of $\mathrm{EtaleDiv}^d(C/k)$ are the effective divisors of degree d which consist of d distinct points, all of which lie in C − Z.]

(9.0.5) Inside $(C^d)_{\text{all dist}}$ we have the open subscheme

$$((C - Z)^d)_{\text{all dist}} \subset (C^d)_{\text{all dist}}$$

whose Y−valued points consist of d−tuples of pairwise disjoint sections Q_i in C(Y), all of which are disjoint from $Z\times_T Y$. We have a cartesian diagram of finite etale S_d−torsors

(9.0.5.1)
$$\begin{array}{ccc} ((C - Z)^d)_{\text{all dist}} & \subset & (C^d)_{\text{all dist}} \\ \downarrow & & \downarrow \\ \mathrm{EtaleDiv}^d(C/k, Z) & \subset & \mathrm{EtaleDiv}^d(C/k). \end{array}$$

(9.0.6) Fix an effective relative Cartier divisor D of degree d in C, and an effective relative Cartier divisor S of C − D of degree $s \geq 0$. We will take the effective relative Cartier divisor Z above to be D + S:

$$Z := D + S.$$

(9.0.7) We have the morphisms

(9.0.7.1)
$$\begin{array}{c} ((C - Z)^d)_{\text{all dist}} \\ \downarrow \\ \mathrm{EtaleDiv}^d(C/T, Z) \\ \downarrow \\ \mathrm{Jac}^d(C/T). \end{array}$$

The divisor class of D is a T−valued point of $\mathrm{Jac}^d(C/T)$, and we view this point as a morphism

(9.0.7.2) $$T \to \mathrm{Jac}^d(C/T).$$

By means of this morphism, we pull back the diagram 9.0.7.1, and obtain a cartesian diagram

(9.0.7.3)
$$\begin{array}{ccc} ((C - Z)^d)_{\text{all dist}, \approx D} & \subset & ((C - Z)^d)_{\text{all dist}} \\ \downarrow & & \downarrow \\ \mathrm{Div}(C, d, D, S) & \subset & \mathrm{EtaleDiv}^d(C/T, Z) \\ \downarrow & & \downarrow \\ T & \to & \mathrm{Jac}^d(C/T) \end{array}$$

which **defines** the closed T−subschemes

$$((C - Z)^d)_{\text{all dist}, \approx D} \quad \subset \quad ((C - Z)^d)_{\text{all dist}}$$

and

$$\mathrm{Div}(C, d, D, S) \quad \subset \quad \mathrm{EtaleDiv}^d(C/T, Z).$$

(9.0.8) Thus Div(C, d, D, S) is the T−scheme whose Y−valued points are the effective relative Cartier divisors of degree d in C which are linearly equivalent to D and which, fppf locally on the base, consist of d distinct points, each of which lies in C − Z.

The top left vertical map in the cartesian diagram 9.0.7.3 above,

$$((C - Z)^d)_{\text{all dist}, \approx D} \to \mathrm{Div}(C, d, D, S),$$

is a finite etale S_d−torsor. The target Div(C, d, D, S) as T−scheme is a fibre−by−fibre open dense set in the projective bundle over T of relative dimension d−g which is the fibre over the class of D in the projective bundle $C^d \to \mathrm{Jac}^d(C/T)$. Thus Div(C, d, D, S) is smooth over T of relative dimension d−g, with geometrically connected fibres. Consequently, $((C - Z)^d)_{\text{all dist}, \approx D}$ is a smooth T−scheme, all of whose fibres are smooth and equidimensional of dimension d − g.

(9.0.8) We have already constructed, in 6.1.10, the T−scheme

$$\mathit{Fct}(C, d, D, S)$$

of functions in L(D) which have d distinct zeroes, all disjoint from Z := D+S. Thus there is a natural map

$$\mathit{Fct}(C, d, D, S) \to \mathrm{Div}(C, d, D, S)$$
$$f \mapsto \text{the divisor of zeroes of } f,$$

which makes *Fct*(C, d, D, S) a Zariski−locally trivial $\mathbb{G}_m$−bundle over Div(C, d, D, S).

(9.0.9) We now return to the finite etale galois S_d−torsor

$$\begin{array}{c} ((C - Z)^d)_{\text{all dist}, \approx D} \\ \downarrow \\ \mathrm{Div}(C, d, D, S). \end{array}$$

We pull back this covering by the natural map

$$\mathit{Fct}(C, d, D, S) \to \mathrm{Div}(C, d, D, S)$$
$$f \mapsto \text{the divisor of zeroes of } f,$$

to get a finite etale galois S_d−torsor

$$\begin{array}{l} \mathrm{Split}(C, d, D, S) := \mathit{Fct}(C, d, D, S) \times_{\mathrm{Div}(C, d, D, S)} ((C - Z)^d)_{\text{all dist}, \approx D} \\ \qquad \downarrow \qquad \swarrow \mathrm{pr}_1 \\ \qquad \mathit{Fct}(C, d, D, S). \end{array}$$

Thus Split(C, d, D, S) is a smooth T−scheme, all of whose fibres are smooth and equidimensional of dimension d + 1 − g.

(9.0.10) The notation Split(C, d, D, S) is inspired by the case when T is the spec of a field k, C is

$\mathbb{P}^1$ and D is d∞. Then a k–valued point f of *Fct*(C, d, D, S) is a polynomial over K of degree d in over variable, say T, with d distinct roots, none of which lies in S. A k–valued point of Split(C, d, D, S) lying over f is an ordered list of d distinct numbers $\alpha_1, ..., \alpha_d$ in k which form a complete factorization, or "splitting" of f, in the sense that

$$f(T) = (\text{elt. of } K^\times) \times \prod_i (T - \alpha_i).$$

Still with T the spec of a field k, in the case of a more general situation (C, D), a k–valued point of Split(C, d, D, S) lying over a k–valued point f of *Fct*(C, d, D, S) is an ordered list of d distinct points $Q_1, ..., Q_d$ in (C – D – S)(k) which form a "splitting" of the divisor of zeroes of f in the sense that

$$\text{div}_0(f) = \Sigma_i\, Q_i$$

(or, equivalently, that div(f) = $\Sigma_i\, Q_i$ – D, but we prefer to focus on the divisor of zeroes of f).

9.1 Theorems of geometric connectedness

Theorem 9.1.1 Hypotheses and notations as in the previous section 9.0, the smooth T–schemes

$$((C - Z)^d)_{\text{all dist}, \approx D} \text{ and Split}(C, d, D, S),$$

which are everywhere of relative dimensions d–g and d+1–g respectively, have geometrically connected (and hence irreducible, because smooth) fibres.

proof Since Split(C, d, D, S) is a Zariski–locally trivial $\mathbb{G}_{m,T}$–bundle over $((C - Z)^d)_{\text{all dist}, \approx D}$, it suffices to show that $((C - Z)^d)_{\text{all dist}, \approx D}$ as T–scheme has geometrically connected fibres. By a standard argument based on the fact that all our data is of finite presentation over T, we reduce to the case when T is affine and of finite type over Spec($\mathbb{Z}$). Covering Spec($\mathbb{Z}$) by Spec($\mathbb{Z}$[1/2]) and Spec($\mathbb{Z}$[1/691]), we may assume further that some prime number ℓ is invertible on T. Denote by

$$\pi: ((C - Z)^d)_{\text{all dist}, \approx D} \to T$$

the structural morphism, and form the sheaf $R^{2(d-g)}\pi_!\overline{\mathbb{Q}}_\ell$ on T. At any geometric point t of T, the dimension of the stalk at t of this sheaf is the number of irreducible components in the fibre $\pi^{-1}(t)$. Thus the set of points of T whose geometric fibre is irreducible is the set of points of T where the stalk of this sheaf is one–dimensional. By the constructibility of $R^{2(d-g)}\pi_!\overline{\mathbb{Q}}_\ell$, it suffices to show that this sheaf has a one–dimensional stalk at every (geometric point over every) closed point. Thus it suffices to treat the case when T is the spec of a finite field k. Since the question is geometric, we may replace k by a finite extension, and suppose further that C(k) is nonempty.

Define

$$h := \dim H_c^{2(d-g)}((((C - D-S)^d)_{\text{all dist}, \approx D})\otimes_k \overline{k}, \overline{\mathbb{Q}}_\ell).$$

Thus h is also the number of connected components of

$$(((C - D-S)^d)_{\text{all dist}, \approx D})^{\otimes_k}\overline{k}.$$

All of these connected components are defined over some finite extension L of k. Over L, each is smooth and geometrically connected, of dimension d−g. So by Lang−Weil, for each finite extension E/L, we have the estimate

$$|\#(((C - D-S)^d)_{\text{all dist}, \approx D})(E) - h(\#E)^{d-g}| = O((\#E)^{d-g-1/2}).$$

Thus to prove the geometric connectedness, we need only prove that for every finite extension E/k, we have an inequality

$$\#(((C - D-S)^d)_{\text{all dist}, \approx D})(E) \le (\#E)^{d-g} + O((\#E)^{d-g-1/2}).$$

To prove this inequality, we consider the morphism of k−schemes

$$C^d \to \mathrm{Jac}^d(C/k),\ (Q_1, ..., Q_d) \mapsto \text{class of } \Sigma_i\, Q_i,$$

and denote by

$$(C^d)_{\approx D} \subset C^d$$

the fibre over the k−valued point class(D) in $\mathrm{Jac}^d(C/k)$. Thus we have an open immersion

$$((C - D-S)^d)_{\text{all dist}, \approx D} \subset (C^d)_{\approx D}.$$

In particular, we have, for every finite extension E/k, an inclusion

$$(((C - D-S)^d)_{\text{all dist}, \approx D})(E) \subset ((C^d)_{\approx D})(E).$$

So it suffices to prove that, for every finite extension E/k, and every E−rational divisor class D of degree d, we have

$$\#((C^d)_{\approx D})(E) = (\#E)^{d-g} + O((\#E)^{d-g-1/2}).$$

This results from the following theorem, applied to $C\otimes_k E/E$.

Theorem 9.1.2 Given integers $g \ge 0$ and $d \ge 2g+1$, there exists an explicit constant

$$\mathrm{Const}(g, d) := 2^d \text{ for } g=0,$$
$$:= (2g-2)^d + (2^{d+2g})\mathrm{Max}(2g, 4), \text{ if } g \ge 1,$$

such that given a finite field k with $\#k \ge 16g^2$, a proper, smooth, geometrically connected curve C/k of genus g with C(k) nonempty, and a divisor D of degree d on C, we have

$$|\#((C^d)_{\approx D})(k) - (\#k)^{d-g}| \le \mathrm{Const}(g,d)(\#k)^{d-g-1/2}.$$

proof Fix a k−rational point P on C. Using P, we get a morphism

$$\pi : C^d \to \mathrm{Jac}^0(C/k),$$
$$\pi(Q_1, ..., Q_d) := \text{class of } \Sigma_i\,(Q_i - P).$$

We also get an isomorphism

$$\mathrm{Jac}^d(C/k) \to \mathrm{Jac}^0(C/k),$$
$$D \mapsto D - dP.$$

So $((C^d)_{\approx D})(k)$ is the set of k−rational points of the fibre of π over the k−rational point D − dP of

$\mathrm{Jac}^0(C/k)$. So we may restate the theorem as

Theorem 9.1.2 bis Given integers $g \geq 0$ and $d \geq 2g+1$, there exists an explicit constant

$$\mathrm{Const}(g, d) := 2^d, \text{ if } g=0,$$
$$= (2g-2)^d + (2^{d+2g})\mathrm{Max}(2g, 4), \text{ if } g \geq 1,$$

with the following property. Given a finite field k with $\#k \geq 16g^2$, a proper, smooth, geometrically connected curve C/k of genus g, and a point P in C(k), form the map

$$\pi : C^d \to \mathrm{Jac}^0(C/k),$$
$$\pi(Q_1, ..., Q_d) := \text{class of } \Sigma_i (Q_i - P).$$

For any divisor class D of degree zero on C, viewed as a k–point of $\mathrm{Jac}^0(C/k)$, we have

$$|\#(\pi^{-1}(D))(k) - (\#k)^{d-g}| \leq \mathrm{Const}(g,d)(\#k)^{d-g-1/2}.$$

proof If g = 0, then C is $\mathbb{P}^1$, $\mathrm{Jac}^0(C/k)$ is a single point, $\#(\pi^{-1}(D))(k)$ is $(\#k + 1)^d$ points, and we may take Const(g,d) to be 2^d.

Suppose now that $g \geq 1$. Let us denote by J/k the Jacobian $\mathrm{Jac}^0(C/k)$, and by F the Frobenius Frob_k. The key idea is to use the Lang torsor

$$1 - F : J \to J,$$

which makes J a finite etale geometrically connected galois covering of itself, with galois group the group J(k) of rational points. Fix a prime ℓ invertible in k. For each $\overline{\mathbb{Q}}_\ell^\times$–valued character ρ of the abelian group J(k), denote by $\mathcal{L}_\rho$ the lisse, rank one, pure of weight zero, $\overline{\mathbb{Q}}_\ell$–sheaf on J obtained from the Lang torsor by extension of the structural group by ρ. At any k–valued point D in J(k), we have

$$\mathrm{Trace}(\mathrm{Frob}_{k,D} \mid \mathcal{L}_\rho) = \rho(D).$$

Moreover, $\mathcal{L}_\rho$ is geometrically nontrivial if and only if ρ is nontrivial.

By orthogonality of characters of finite abelian groups, the characteristic function I_D of an element D in J(k) is given by the sum

$$I_D = (1/\#J(k))\Sigma_\rho \overline{\rho}(D)\rho.$$

Therefore we have

$$\#(\pi^{-1}(D))(k) = (1/\#J(k))\Sigma_\rho \overline{\rho}(D)\Sigma_{(Q_1, ..., Q_d) \text{ in } C^d(k)} \rho(\Sigma_i (Q_i - P)).$$

We move the term corresponding to the trivial character to the other side of this equality to obtain

$$\#(\pi^{-1}(D))(k) - (\#C(k))^d/\#J(k)$$
$$= (1/\#J(k))\Sigma_{\rho \text{ nontriv}} \overline{\rho}(D)\Sigma_{(Q_1, ..., Q_d) \text{ in } C^d(k)} \rho(\Sigma_i (Q_i - P)).$$

At this point we need the following fundamental estimate:

Proposition 9.1.3 Notations as in 9.1.2, if ρ is a nontrivial character of J(k) we have the estimate

$$|\Sigma_{(Q_1, ..., Q_d) \text{ in } C^d(k)}\, \rho(\Sigma_i (Q_i - P))| \le (2g-2)^d(\#k)^{d/2}.$$

proof The sum in question is the d'th power of the sum for d=1:

$$\Sigma_{(Q_1, ..., Q_d) \text{ in } C^d(k)}\, \rho(\Sigma_i (Q_i - P)) = (\Sigma_{Q \text{ in } C(k)}\, \rho(Q - P))^d.$$

So what we must prove is the estimate

$$|\Sigma_{Q \text{ in } C(k)}\, \rho(Q - P)| \le (2g-2)(\#k)^{1/2}.$$

Let us denote by $\varphi : C \to J$ the embedding $\varphi(Q) :=$ class of $Q - P$. Then φ induces an isomorphism of abelianized geometric fundamental groups

$$(\varphi_*)^{ab} : \pi_1(C\otimes_k\bar{k})^{ab} \cong \pi_1(J\otimes_k\bar{k})^{ab} = \pi_1(J\otimes_k\bar{k}).$$

Therefore $\varphi^*\mathcal{L}_\rho$ is geometrically nontrivial on C. As $\varphi^*\mathcal{L}_\rho$ is lisse of rank one on C, we have

$$\chi(C\otimes_k\bar{k}, \varphi^*\mathcal{L}_\rho) = \chi(C\otimes_k\bar{k}, \bar{\mathbb{Q}}_\ell) = 2-2g.$$

Because $\varphi^*\mathcal{L}_\rho$ is geometrically nontrivial on C and lisse of rank one, we have

$$H^0(C\otimes_k\bar{k}, \varphi^*\mathcal{L}_\rho) = 0 = H^2(C\otimes_k\bar{k}, \varphi^*\mathcal{L}_\rho).$$

Thus $h^1(C\otimes_k\bar{k}, \varphi^*\mathcal{L}_\rho) = 2g-2$, and, by Deligne, $H^1(C\otimes_k\bar{k}, \varphi^*\mathcal{L}_\rho)$ is pure of weight one.

The Lefschetz Trace Formula gives

$$\Sigma_{Q \text{ in } C(k)}\, \rho(Q - P) = -\mathrm{Trace}(\mathrm{Frob}_k \mid H^1(C\otimes_k\bar{k}, \varphi^*\mathcal{L}_\rho)),$$

so we get the required estimate

$$|\Sigma_{Q \text{ in } C(k)}\, \rho(Q - P)| \le (2g-2)(\#k)^{1/2}. \qquad \text{QED}$$

We now conclude the proof of Theorem 9.1.2bis. Using this estimate for each of the (#J(k) − 1) nontrivial characters ρ, we get

(9.1.3.1) $$|\#(\pi^{-1}(D))(k) - (\#C(k))^d/\#J(k)| \le (2g-2)^d(\#k)^{d/2}.$$

By Weil, we have

$$(1 - 2g(\#k)^{-1/2})^d \le (\#C(k)/\#k)^d \le (1 + 2g(\#k)^{-1/2})^d,$$

and

$$(1 + (\#k)^{-1/2})^{2g} \le \#J(k)/(\#k)^g \le (1 - (\#k)^{-1/2})^{2g}.$$

Thus

$$(\#C(k))^d/\#J(k) \ge (\#k)^{d-g}(1 - 2g(\#k)^{-1/2})^d/(1 + (\#k)^{-1/2})^{2g}.$$
$$\ge (\#k)^{d-g}(1 - 2g(\#k)^{-1/2})^d(1 - (\#k)^{-1/2})^{2g}$$

(using the inequality $1/(1+x) \ge 1 - x$ for real x in [0, 1])

and

$$(\#C(k))^d/\#J(k) \le (\#k)^{d-g}(1 + 2g(\#k)^{-1/2})^d/(1 - (\#k)^{-1/2})^{2g}$$

$$\leq (\#k)^{d-g}(1 + 2g(\#k)^{-1/2})^d(1 + 4(\#k)^{-1/2})^{2g},$$

(using the inequality $1/(1-x) \leq 1+4x$ for real x in $[0, 1/\mathrm{Sqrt}(2)]$).

These inequalities in turn imply

$$(\#C(k))^d/\#J(k) \geq (\#k)^{d-g}(1 - 2g(\#k)^{-1/2})^{d+2g},$$

and

$$(\#C(k))^d/\#J(k) \leq (\#k)^{d-g}(1 + \mathrm{Max}(2g,4)(\#k)^{-1/2})^{d+2g}.$$

For real x in $[0, 1]$, and any integer $n \geq 1$, we have the inequality

$$(1 - x)^n \geq 1 - (2^n - 1)x \geq 1 - 2^n x.$$

Since $\#k \geq 16g^2$, we may apply this with $x = 2g(\#k)^{-1/2}$, and we find

$$(\#C(k))^d/\#J(k) \geq (\#k)^{d-g}(1 - 2g(\#k)^{-1/2})^{d+2g}$$
$$\geq (\#k)^{d-g}(1 - (2^{d+2g})2g(\#k)^{-1/2}).$$

For real x in $[0, 1]$, and any integer $n \geq 1$, we have the inequality

$$(1 + x)^n \leq 1 + (2^n - 1)x \leq 1 + 2^n x.$$

Since $\#k \geq 16g^2$ and $g \geq 1$, we may apply this with $x = \mathrm{Max}(2g,4)(\#k)^{-1/2}$, and we find

$$(\#C(k))^d/\#J(k) \leq (\#k)^{d-g}(1 + \mathrm{Max}(2g,4)(\#k)^{-1/2})^{d+2g}$$
$$\leq (\#k)^{d-g}(1 + (2^{d+2g})\mathrm{Max}(2g,4)(\#k)^{-1/2}).$$

Thus we have

$$|(\#C(k))^d/\#J(k) - (\#k)^{d-g}| \leq (2^{d+2g})\mathrm{Max}(2g,4)(\#k)^{d-g-1/2}.$$

Combining this with the previous estimate (9.1.3.1),

$$|\#(\pi^{-1}(D))(k) - (\#C(k))^d/\#J(k)| \leq (2g-2)^d(\#k)^{d/2},$$

we get

$$|\#(\pi^{-1}(D))(k) - (\#k)^{d-g}| \leq (2g-2)^d(\#k)^{d/2} + (2^{d+2g})\mathrm{Max}(2g,4)(\#k)^{d-g-1/2}.$$

But $d \geq 2g+1$, so $d/2 \leq d - g - 1/2$, so we have

$$|\#(\pi^{-1}(D))(k) - (\#k)^{d-g}| \leq \mathrm{Const}(g, d)(\#k)^{d-g-1/2},$$

with

$$\mathrm{Const}(g, d) := (2g-2)^d + (2^{d+2g})\mathrm{Max}(2g,4).$$ QED for 9.1.2bis

Corollary 9.1.4 Let k be a finite field, C/k a proper, smooth, geometrically connected curve of genus g. For any integer $d \geq 2g+1$, the natural map

$$C^d \to \mathrm{Jac}^d(C/k)$$

has geometrically irreducible fibres.

proof The morphism $\pi : C^d \to \mathrm{Jac}^d(C/k)$ is flat, being the composition of the finite flat map $C^d \to \mathrm{Sym}^d(C/k)$ with the projective bundle $\mathrm{Sym}^d(C/k) \to \mathrm{Jac}^d(C/k)$. Both source and target of π are smooth and equidimensional, of dimensions d and g respectively. So every fibre of π is a local complete intersection, equidimensional of dimension $d-g$. Therefore our diophantine estimate

9.1.2, together with Lang–Weil, shows that π has geometrically irreducible fibres. QED

9.2 Interpretation in terms of geometric monodromy groups

Theorem 9.2.1 Hypotheses and notations as in 9.0, suppose further that T is connected. Consider the finite etale S_d–torsor

$$\begin{array}{c}\mathrm{Split}(C, d, D, S)\\ \downarrow\\ \mathit{Fct}(C, d, D, S).\end{array}$$

1) For any geometric point t of T, and any geometric point ξ_t of

$$\mathit{Fct}(C, d, D, S)t = \mathit{Fct}(C_t, d, D_t, S_t),$$

the classifying map

$$\rho_{\mathrm{split},t} : \pi_1(\mathit{Fct}(C, d, D, S)_t, \xi_t) \to S_d$$

for the pullback S_d torsor

$$\begin{array}{c}\mathrm{Split}(C_t, d, D_t, S_t)\\ \downarrow\\ \mathit{Fct}(C_t, d, D_t, S_t)\end{array}$$

is surjective.

2) For any geometric point ξ of $\mathit{Fct}(C, d, D, S)$, the corresponding group homomorphism

$$\rho_{\mathrm{split}} : \pi_1(\mathit{Fct}(C, d, D, S), \xi) \to S_d$$

which "classifies" this finite etale S_d–torsor is surjective.

proof 1) The surjectivity is equivalent to the connectedness of the total space $\mathrm{Split}(C_t, d, D_t, S_t)$. This connectedness is proven in Theorem 9.1.1 above. Assertion 2) is a formal consequence of 1). Indeed, the question is independent of the choice of the base point ξ, which we will now choose conveniently. Pick a geometric point t of T, and a geometric point ξ_t of $\mathit{Fct}(C, d, D, S)t = \mathit{Fct}(C_t, d, D_t, S_t)$. Then $\rho_{\mathrm{split},t}$ is the composite group homomorphism

$$\pi_1(\mathit{Fct}(C, d, D, S)_t, \xi_t) \xrightarrow{\mathrm{inclusion}_*} \pi_1(\mathit{Fct}(C, d, D, S), \xi_t) \xrightarrow{\rho_{\mathrm{split}}} S_d.$$

As the composite $\rho_{\mathrm{split},t}$ is surjective by part 1), ρ_{split} itself must be surjective. QED

(9.2.2) We now wish to translate the above result into one about geometric monodromy groups of lisse sheaves. To do this in as straightforward a way as possible, for each integer $d \geq 1$, denote by π_d the d–dimensional representation of S_d on linear forms in d variables,

$$\pi_d : S_d \to O(d, \mathbb{Z}) \subset GL(d, \mathbb{Z}).$$

We can push out the S_d–torsor

$$\begin{array}{c} \mathrm{Split}(C, d, D, S) \\ \downarrow \\ \mathit{Fct}(C, d, D, S). \end{array}$$

by π_d, and we obtain on the space $\mathit{Fct}(C, d, D, S)$ a sheaf $\mathcal{S}_d$ of free $\mathbb{Z}$–modules of rank d which is literally locally constant in the etale topology. For any prime number ℓ, we can form

$$\mathcal{S}_{d,\ell} := \mathcal{S}_d \otimes_{\mathbb{Z}} \overline{\mathbb{Q}}_\ell,$$

which is now a lisse $\overline{\mathbb{Q}}_\ell$–sheaf on $\mathit{Fct}(C, d, D, S)$ which is literally locally constant in the etale topology on $\mathit{Fct}(C, d, D, S)$. It is ι–pure of weight zero for every ι, since every eigenvalue of every Frobenius is a root of unity of order dividing d!.

Corollary 9.2.3 Hypotheses as in Theorem 9.2.1, suppose in addition that T is a normal connected scheme which is of finite type over $\mathbb{Z}[1/\ell]$ for some prime ℓ. Denote by X the space

$$X := \mathit{Fct}(C, d, D, S).$$

Thus X/T is smooth of relative dimension d+1−g, with geometrically connected fibres. Consider the lisse, rank d $\overline{\mathbb{Q}}_\ell$–sheaf

$$\mathcal{S} := \mathcal{S}_{d,\ell}$$

on X. Denote by η the generic point of T, by $\overline{\eta}$ a geometric generic point of T, and by ξ a geometric point of $X_{\overline{\eta}}$. Denote by

$$\rho_{\mathcal{S}} : \pi_1(X, \xi) \to S_d \subset GL(d, \overline{\mathbb{Q}}_\ell)$$

the representation of $\pi_1(X, \xi)$ which $\mathcal{S}$ "is". For every finite field k, and every k–valued point t of T, the group G_{geom} for $\mathcal{S}_t$:= the restriction of $\mathcal{S}$ to X_t/k is (conjugate in $GL(d, \overline{\mathbb{Q}}_\ell)$ to) S_d.

proof This is the special case of Theorem 9.2.1 in which T is the spec of a finite field. QED

9.3 Relation to "splitting of primes"

(9.3.1) Let k be a finite field, and t a k–valued point of T. Given a finite extension E/k, and an E–valued point of X_t

$$f \text{ in } X_t(E) := \mathit{Fct}(C_t, d, D_t, S_t)(E),$$

its Frobenius conjugacy class

$$\rho_{split}(\mathrm{Frob}_{E,f}) \text{ in } (S_d)^{\#}$$

has a straightforward description in terms of how the divisor of zeroes of f, $\mathrm{div}_0(f)$, "factors" over E. We are given that, over $\overline{E}$, $\mathrm{div}_0(f)$ consists of d distinct points in $C_t(\overline{E})$. Break the set of these points into orbits under $\mathrm{Gal}(\overline{E}/E)$, i.e., write $\mathrm{div}_0(E)$ as a sum of distinct closed points of $C_t \otimes_k E$, say

$$\mathrm{div}_0(f) = \Sigma_i \mathcal{P}_i.$$

The degrees n_i of the closed points $\mathcal{P}_i$ are the cardinalities of the orbits of $\mathrm{Gal}(\bar{E}/E)$ acting on $\mathrm{div}_0(f)(\bar{E})$. These degrees n_i form an unordered partition of d. The Frobenius conjugacy class

$$\rho_{\mathrm{split}}(\mathrm{Frob}_{E,f}) \text{ in } (S_d)^\#$$

is the conjugacy class named by this partition of d, namely the conjugacy class of a product of disjoint cycles of lengths the n_i.

(9.3.2) We say that f is a **prime** in $X_t(E)$ if its Frobenius conjugacy class is a d–cycle, or, equivalently, if its divisor of zeroes is a single closed point in $C_t \otimes_k E$ (necessarily of degree d). [For example, in the case when C_t is $\mathbb{P}^1$ and D is d∞, a prime f in $X_t(E)$ is precisely an irreducible polynomial of degree d in E[T] which is invertible on S.] We denote by

$$X_{t,\mathrm{prime}}(E) \subset X_t(E)$$

the set of primes in $X_t(E)$.

(9.3.3) More generally, for any conjugacy class (:= partition of d) σ in S_d, we say that f in $X_t(E)$ is of splitting type σ if its Frobenius conjugacy class $\rho_{\mathrm{split}}(\mathrm{Frob}_{E,f})$ in $(S_d)^\#$ is in the class σ. We denote by

$$X_{t,\sigma-\mathrm{split}}(E) \subset X_t(E)$$

the set of elements of $X_t(E)$ of splitting type σ.

(9.3.4) So in this somewhat cumbersome terminology, a prime in $X_t(E)$ is an element of splitting type σ for σ the class of a d–cycle (the partition d=d of d). At the other extreme, if we take for σ the conjugacy class {e}, corresponding to the partition $d = \Sigma\, 1$, we get the notion of a totally split f in $X_t(E)$, a function whose zeroes are d distinct E–rational points.

9.4 Distribution of primes in the spaces $X_t := \mathit{Fct}(C_t, d, D_t, S_t)$

(9.4.1) Before stating the main result 9.4.4 of this section, we must recall two definitions [Ka–Sar, RMFEM, 9.2.6, 5) and 4)]. We fix a prime number ℓ. Given an algebraically closed field k in which ℓ is invertible, and X/k a smooth connected k–scheme of dimension d, we define the nonnegative integer A(X) by

$$A(X) := \Sigma_{i<2d}\, h_c{}^i(X, \bar{\mathbb{Q}}_\ell).$$

Given a lisse $\bar{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on X, we define the nonnegative integer $C(X, \mathcal{F})$ as follows. There exists a finite extension E_λ of $\mathbb{Q}_\ell$ with integer ring $\mathcal{O}_\lambda$ and residue field $\mathbb{F}_\lambda$, a lisse torsion–free $\mathcal{O}_\lambda$–form $\mathcal{F}_{\mathcal{O}_\lambda}$ of $\mathcal{F}$, and a finite etale $\pi : Y \to X$, Y not necessarily connected, such that $\mathcal{F}_{\mathcal{O}_\lambda} \otimes_{\mathcal{O}_\lambda} \mathbb{F}_\lambda$ becomes trivial after pullback to Y. For each choice $(E_\lambda, \mathcal{F}_{\mathcal{O}_\lambda}, Y)$ of such data, we define

$$C(X, \mathcal{F}, E_\lambda, \mathcal{F}_{\mathcal{O}_\lambda}, Y) := \sum_i h_c{}^i(Y, \mathbb{F}_\lambda).$$

We define $C(X, \mathcal{F})$ to be the minimum value of $C(X, \mathcal{F}, E_\lambda, \mathcal{F}_{\mathcal{O}_\lambda}, Y)$ over all choices of $(E_\lambda, \mathcal{F}_{\mathcal{O}_\lambda}, Y)$.

(9.4.2) Both of these quantities remain bounded when the data moves in a family.

Uniformity Lemma 9.4.3 Let T be a normal connected $\mathbb{Z}[1/\ell]$–scheme of finite type, X/T a smooth T–scheme with geometrically connected fibres of dimension d, $\mathcal{F}$ a lisse $\overline{\mathbb{Q}}_\ell$–sheaf on X. There exist nonnegative integers A(X/T) and C(X/T, $\mathcal{F}$) such that for every geometric point t of T, we have

$$A(X_t) \le A(X/T),$$
$$C(X_t, \mathcal{F}|X_t) \le C(X/T, \mathcal{F}).$$

proof See [Ka–Sar, RMFEM, 9.3.3 and 9.3.4]. QED

Theorem 9.4.4 Hypotheses and notations as in Corollary 9.2.3, we have the following results. For any finite field k with $\mathrm{Card}(k) \ge 4A(X/T)^2$, any conjugacy class σ in S_d, any k–valued point t of T, and any finite extension E/k, we have

$$|\#X_{t,\sigma\text{-split}}(E)/\#X_t(E) - \#\sigma/d!| \le 2C(X/T, \mathcal{S})d!/(\#E)^{1/2}.$$

In particular, taking σ to be the class of a d–cycle, we have

$$|\#X_{t,\mathrm{prime}}(E)/\#X_t(E) - 1/d| \le 2C(X/T, \mathcal{S})d!/(\#E)^{1/2}.$$

proof Apply Deligne's equidistribution theorem [Ka–Sar, RMFEM, 9.7.13], with the data

$$(\ell, X/S, \mathcal{F}, \iota, G, G_{\mathrm{arith}})$$

of [Ka–Sar, RMFEM, 9.7.10] taken to be

$$(\ell, X/T, \mathcal{S}, \text{any } \iota, G = G_{\mathrm{arith}} = S_d \text{ inside } GL(d)).$$

In the notations of [Ka–Sar, RMFEM, 9.7.13], $K = K_{\mathrm{arith}} = S_d$, γ is the unique element of the group $\Gamma = \{e\}$, and we take for W the conjugacy class σ. We have already observed that $\mathcal{S}$ is ι–pure of weight zero. That the other hypotheses [Ka–Sar, RMFEM, 9.7.2.1–3] hold is precisely the content of Corollary 9.2.3 above. QED

9.5 Equidistribution theorems for twists by primes: the basic setup over a finite field

(9.5.1) In order to clarify the simple underlying structure, we will first consider a slightly simplified abstract situation.We give ourselves a finite field k, a smooth, geometrically connected k–scheme X/k, a geometric point ξ of X, a prime number ℓ invertible in k, a field embedding $\iota : \overline{\mathbb{Q}}_\ell \to \mathbb{C}$, and a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on X of rank r, which is ι–pure of weight zero. We denote by

$$\Theta_{\mathcal{F}}: \pi_1(X, \xi) \to GL(\mathcal{F}_\xi) \cong GL(r, \overline{\mathbb{Q}}_\ell)$$

the homomorphism corresponding to the lisse sheaf $\mathcal{F}$. We denote by $G_{geom,\mathcal{F}}$ the Zariski closure in $GL(\mathcal{F}_\xi)$ of the image of $\pi_1{}^{geom}(X, \xi) := \pi_1(X\otimes_k\overline{k}, \xi)$ under $\Theta_{\mathcal{F}}$. We denote by $G_{arith,\mathcal{F}}$ the Zariski closure in $GL(\mathcal{F}_\xi)$ of the image of $\pi_1{}^{arith}(X, \xi) := \pi_1(X, \xi)$ under $\Theta_{\mathcal{F}}$. Thus $G_{geom,\mathcal{F}}$ is a closed normal subgroup of $G_{arith,\mathcal{F}}$.

(9.5.2) We make the hypothesis that $G_{geom,\mathcal{F}}$ is of finite index in $G_{arith,\mathcal{F}}$, and we denote by S the finite quotient group:

$$S := G_{arith.\mathcal{F}}/G_{geom,\mathcal{F}}.$$

The group S is a finite cyclic group, because it is a finite quotient of the procyclic group $\pi_1{}^{arith}(X, \xi)/\pi_1{}^{geom}(X, \xi) \cong Gal(\overline{k}/k)$. Thus S has a canonical generator, the image of the geometric Frobenius $Frob_k$ in $Gal(\overline{k}/k)$. Thus $S = \mathbb{Z}/(\#S)\mathbb{Z}$. We will speak of elements of S as "degrees mod #S".

(9.5.3) We pick maximal compact subgroups K of $G_{geom,\mathcal{F}}(\mathbb{C})$ and K_{arith} of $G_{arith,\mathcal{F}}(\mathbb{C})$ with $K \subset K_{arith}$. Then $K_{arith}/K \cong S$. We denote by dk the Haar measure on K_{arith} which gives K total mass one (and so gives K_{arith} total mass #S). For each s in S, we denote by $K_{arith,s} \subset K_{arith}$ the coset sK. The surjective homomorphism

$$K_{arith} \to S$$

induces a map of spaces of conjugacy classes

$$(K_{arith})^\# \to S^\# = S.$$

For each s in S, we denote by $(K_{arith,s})^\# \subset (K_{arith})^\#$ the inverse image of s by this map.

(9.5.4) Fix one element s in S. For E/k any finite extension whose degree is congruent to s mod #S, and any x in X(E), the element $\iota(\Theta_{\mathcal{F}}(Frob_{E,x}))^{ss}$ in $G_{arith,\mathcal{F}}(\mathbb{C})$ is conjugate in $G_{arith,\mathcal{F}}(\mathbb{C})$ to an element $\theta(E, x)$ of $K_{arith,s}$, and this element $\theta(E, x)$ is itself well defined up to K_{arith}–conjugacy. By Deligne's equidistribution theorem [Ka–Sar, RMFEM, 9.7.10], we know that for any continuous $\mathbb{C}$–valued central function f on K_{arith}, we have the limit formula

(9.5.4.1) $$\int_{K_{arith,s}} f(k)dk = \lim_{\#E \to \infty,\ deg(E/k) \equiv s \bmod \#S} (1/\#X(E))\Sigma_{x \text{ in } X(E)} f(\theta(E, x)),$$

the limit taken over finite extensions E/k of degree ≡ s mod #S and large enough that X(E) is nonempty. More precisely, for Λ any finite–dimensional representation of K_{arith}, and any finite extension E/k of degree ≡ s mod #S with $Card(E) \geq 4A(X\otimes_k\overline{k})^2$, we have the estimate

(9.5.4.2)

$$\left|\int_{K_{arith,s}} Trace(\Lambda(k))dk - (1/\#X(E))\Sigma_{x \text{ in } X(E)} Trace(\Lambda(\theta(E, x)))\right|$$
$$\leq 2C(X\otimes_k\overline{k}, \mathcal{F})\dim(\Lambda)/Card(E)^{1/2}.$$

(9.5.5) We also give ourselves a **finite** group Γ, and a homomorphism

$$\rho : \pi_1(X, \xi) \to \Gamma.$$

We suppose that

$$\rho(\pi_1{}^{geom}(X, \xi)) = \Gamma.$$

We choose a faithful $\overline{\mathbb{Q}}_\ell$ representation of the finite group Γ, and view it as a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{S}_\Gamma$ on X which becomes trivial on a finite etale covering (the one determined by Ker(ρ)).

(9.5.6) For each conjugacy class γ in Γ, and each finite extension E/k, we denote by

$$X_\gamma(E) \subset X(E)$$

the set of points x in X(E) such that the Frobenius conjugacy class $\rho(\mathrm{Frob}_{E,x})$ lies in the class γ.

(9.5.7) Applying [Ka–Sar, RMFEM, 9.7.2.13], we find that for any finite extension E/k with $\mathrm{Card}(E) \geq 4A(X\otimes_k\overline{k})^2$, and any conjugacy class γ in Γ, we have

(9.5.7.1)) $$|\#X_\gamma(E)/\#X(E) - \#\gamma/\#\Gamma| \leq 2C(X\otimes_k\overline{k}, \mathcal{S}_\Gamma)\#\Gamma/(\#E)^{1/2}.$$

Lemma 9.5.8 For $\mathrm{Card}(E) > \mathrm{Max}(4A(X\otimes_k\overline{k})^2, 4C(X\otimes_k\overline{k}, \mathcal{S}_\Gamma)^2(\#\Gamma)^4)$, both X(E) and $X_\gamma(E)$ are nonempty.

proof We recall that for $\mathrm{Card}(E) > 4A(X\otimes_k\overline{k})^2$, we have $\mathrm{Card}(X(E)) \geq (1/2)\mathrm{Card}(E)^{\dim(X)}$, so certainly X(E) is nonempty. Thus we will have $X_\gamma(E)$ nonempty provided $2C(X\otimes_k\overline{k}, \mathcal{S}_\Gamma)\#\Gamma/(\#E)^{1/2} < 1/\#\Gamma$, or, what is the same, provided that

$$\mathrm{Card}(E) > 4C(X\otimes_k\overline{k}, \mathcal{S}_\Gamma)^2(\#\Gamma)^4. \qquad \text{QED}$$

(9.5.9) Now let us return our attention to Deligne's equidistribution theorem for $\mathcal{F}$:

$$\int_{K_{arith,s}} f(k)dk = \lim_{\#E \to \infty,\ \deg(E/k) \equiv s \bmod \#S} (1/\#X(E))\Sigma_{x \text{ in } X(E)} f(\theta(E, x)),$$

the limit taken over finite extensions E/k of degree ≡ s mod #S and large enough that X(E) is nonempty. Fix a conjugacy class γ in Γ. We are interested in the extent to which this formula remains true if we replace, in its right hand side, the average over X(E) by the average over $X_\gamma(E)$. In other words, when is it true that

$$\int_{K_{arith,s}} f(k)dk = \lim_{\#E \to \infty,\ \deg(E/k) \equiv s \bmod \#S} (1/\#X_\gamma(E))\Sigma_{x \text{ in } X_\gamma(E)} f(\theta(E, x)),$$

the limit now taken over finite extensions E/k of degree ≡ s mod #S large enough that $X_\gamma(E)$ is nonempty?

(9.5.10) To answer this question, we must consider the homomorphism

$$\Theta_{\mathcal{F}}\times\rho : \pi_1(X, \xi) \to G_{geom,\mathcal{F}}(\overline{\mathbb{Q}}_\ell) \times \Gamma.$$

Denote by $G_{geom,\mathcal{F}\times\Gamma}$ the Zariski closure of $(\Theta_{\mathcal{F}}\times\rho)(\pi_1{}^{geom}(X, \xi))$ in $G_{geom,\mathcal{F}}\times\Gamma$. Denote by $G_{arith,\mathcal{F}\times\Gamma}$ the Zariski closure of $(\Theta_{\mathcal{F}}\times\rho)(\pi_1{}^{arith}(X, \xi))$ in $G_{arith,\mathcal{F}}\times\Gamma$.

Theorem 9.5.11 Suppose the group $G_{geom,\mathcal{F}\times\Gamma}$ is equal to the product $G_{geom,\mathcal{F}}\times \Gamma$. Then $G_{arith,\mathcal{F}\times\Gamma}$ is equal to the product $G_{arith,\mathcal{F}}\times \Gamma$. For every s in S, every conjugacy class γ in Γ, and every continuous $\mathbb{C}$–valued central function f on K_{arith}, we have the limit formula

$$\int_{K_{arith,s}} f(k)dk = \lim_{\#E \to \infty,\ deg(E/k) \equiv s \bmod \#S} (1/\#X_\gamma(E))\Sigma_{x \text{ in } X_\gamma(E)} f(\theta(E, x)),$$

the limit taken over finite extensions E/k of degree $\equiv$ s mod #S large enough that $X_\gamma(E)$ is nonempty, e.g, Card(E) > $\mathrm{Max}(4A(X\otimes_k\bar{k})^2, 4C(X\otimes_k\bar{k}, \mathcal{S}_\Gamma)^2(\#\Gamma)^4)$.

proof The group G_{arith} for $\mathcal{F}\times\Gamma$ lies in the product $G_{arith,\mathcal{F}}\times \Gamma$ and projects onto each factor. It contains as a subgroup the group G_{geom} for $\mathcal{F}\times\Gamma$, which by hypothesis is the product $G_{geom,\mathcal{F}}\times \Gamma$. Thus $G_{arith,\mathcal{F}\times\Gamma}$ is a group between $G_{geom,\mathcal{F}}\times \Gamma$ and $G_{arith,\mathcal{F}}\times \Gamma$ which maps onto $G_{arith,\mathcal{F}}$, so must be the product $G_{arith,\mathcal{F}}\times \Gamma$.

Pick any faithful linear $\bar{\mathbb{Q}}_\ell$–representation Λ of Γ, say of dimension n, and denote by $\mathcal{S}_\Gamma$ the lisse $\bar{\mathbb{Q}}_\ell$–sheaf on X of rank n attached to the composite homomorphism

$$\Lambda\circ\rho : \pi_1(X) \to \Gamma \to GL(n, \bar{\mathbb{Q}}_\ell).$$

We now apply Deligne's equidistribution theorem as recalled above to the direct sum sheaf $\mathcal{F}\oplus\mathcal{S}_\Gamma$. The group $G_{geom,\mathcal{F}\oplus\mathcal{S}_\Gamma}$ is, by hypothesis, the product group $G_{geom,\mathcal{F}}\times \Gamma$. As we have just seen above, $G_{arith,\mathcal{F}\oplus\mathcal{S}_\Gamma}$ is the product group $G_{arith,\mathcal{F}}\times \Gamma$. A maximal compact subgroup of $G_{geom,\mathcal{F}}\times \Gamma$ is $K\times\Gamma$, and a maximal compact subgroup of $G_{arith,\mathcal{F}\oplus\mathcal{S}_\Gamma}$ is $K_{arith}\times\Gamma$.

Fix a conjugacy class γ in Γ, and denote by

$$I_\gamma : \Gamma \to \mathbb{C}$$

the indicator function of the conjugacy class γ. Denote by dg the total mass one Haar measure on Γ. Given a continuous $\mathbb{C}$–valued central function f on K_{arith}, the product function $f\times I_\gamma$ on $K_{arith}\times\Gamma$ is a continuous $\mathbb{C}$–valued central function. For each s in S, Deligne's equidistribution theorem [Ka–Sar, RMFEM, 9.7.10] for $\mathcal{F}\oplus\mathcal{S}_\Gamma$ gives

$$\int_{K_{arith,s}\times\Gamma} f(k)I_\gamma(g)dkdg$$
$$= \lim_{\#E \to \infty,\ deg(E/k) \equiv s \ mo}(1/\#X(E))\Sigma_{x \text{ in } X(E)} f(\theta(E, x))I_\gamma(\rho(\mathrm{Frob}_{E,x})),$$

the limit taken over finite extensions E/k of degree $\equiv$ s mod #S large enough that X(E) is nonempty. More explicitly, this limit formula says

$$(\#\gamma/\#\Gamma)\int_{K_{arith,s}} f(k)dk = \lim_{\#E \to \infty,\ deg(E/k) \equiv s \bmod \#S} (1/\#X(E))\Sigma_{x \text{ in } X_\gamma(E)} f(\theta(E, x)),$$

the limit taken over finite extensions E/k of degree $\equiv$ s mod #S large enough that X(E) is nonempty. We also know from 9.5.7.1 that

$$(\#\Gamma/\#\gamma) = \lim_{\#E \to \infty} (\#X(E)/\#X_\gamma(E)),$$

the limit taken over finite extensions E/k large enough that $X_\gamma(E)$ is nonempty. In particular,we have

$$(\#\Gamma/\#\gamma) = \lim_{\#E \to \infty} (\#X(E)/\#X_\gamma(E)),$$

the limit taken over finite extensions E/k of degree ≡ s mod #S large enough that X(E) is nonempty. Multiplying together these two limit formulas, we get the assertion. QED

9.6 Equidistribution theorems for twists by primes: uniformities with respect to parameters in the basic setup above

(9.6.1) In this section, we consider the following situation. We are given a prime number ℓ, a field embedding $\iota : \overline{\mathbb{Q}}_\ell \to \mathbb{C}$, a connected normal $\mathbb{Z}[1/\ell]$–scheme T of finite type, a smooth T–scheme X/T with geometrically connected fibres of dimension d, a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on X of rank r ≥ 1, a finite group Γ, and a finite etale galois Γ–torsor Y/X on X. We choose a faithful $\overline{\mathbb{Q}}_\ell$–linear representation of Γ, and push out Y/X by this representation to obtain a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{S}_\Gamma$ on X which becomes trivial on Y. We fix two (not necessarily connected) semisimple $\overline{\mathbb{Q}}_\ell$–algebraic subgroups $G \subset G_{arith}$ of GL(r). We suppose that G is a normal subgroup of G_{arith} of finite index, and that the quotient group G_{arith}/G is a finite cyclic group S. We fix maximal compact subgroups K in $G(\mathbb{C})$ and K_{arith} in $G_{arith}(\mathbb{C})$, with $K \subset K_{arith}$. We make the following hypothesis:

(9.6.2) For every finite field k, and every k–valued point t of T, there exists a constant $\alpha_{k,t}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ such that

1) The lisse sheaf on X_t/k given by $\mathcal{F}_t \otimes (\alpha_{k,t})^{deg}$ is ι–pure of weight zero.

2) The group G_{geom} for $\mathcal{F}_t \otimes (\alpha_{k,t})^{deg}$ is (conjugate in GL(r) to) to G.

3) Under the representation ρ_t of $\pi_1(X_t)$ corresponding to $\mathcal{F}_t \otimes (\alpha_{k,t})^{deg}$, the entire group $\pi_1(X_t)$ lands in G_{arith}, i.e., we have $\rho_t(\pi_1(X_t)) \subset G_{arith}$.

4) The group G_{geom} for the direct sum $\mathcal{F}_t \otimes (\alpha_{k,t})^{deg} \oplus \mathcal{S}_{\Gamma,t}$ on X_t is the product group (G_{geom} for $\mathcal{F}_t \otimes (\alpha_{k,t})^{deg}$)×Γ.

5) There exists a surjective homomorphism

$$a : \pi_1(T) \to S$$

with the following property. For each finite field k, each k–valued point t in T(k), and each k–valued point x in $X_t(k)$, the image in S of $\rho_t(\mathrm{Frob}_{k,x})$ is $A(\mathrm{Frob}_{k,t})$.

Theorem 9.6.3 Notations and hypotheses as in 9.6.1–2 above, fix an element s in S, and a conjugacy class γ in Γ. For each finite field k and each k–valued point t of T such that $A(\mathrm{Frob}_{k,t}) = s$, and each k–valued point x of X_t with Frobenius conjugacy class γ in Γ, denote by $\theta(k, t, \alpha_{k,t}, x)$ the Frobenius conjugacy class in $K_{arith,s}$ attached to the point x and the lisse sheaf

$\mathcal{F}_t\otimes(\alpha_{k,t})^{\deg}$ on X_t. Fix a continuous $\mathbb{C}$-valued central function f on K_{arith}. Fix any sequence of data (k_i, t_i in $T(k_i)$) in which the k_i are finite fields with

$$\mathrm{Card}(k_i) > \mathrm{Max}(4A(X/T)^2, 4C(X/T, \mathcal{S}_\Gamma)^2(\#\Gamma)^4)$$

whose cardinalities form a strictly increasing sequence, and in which, for each i, $A(\mathrm{Frob}_{k_i,t_i}) = s$. We have the limit formula

$$\int_{K_{arith,s}} f(k)dk = \lim_{i \to \infty} (1/\#X_{t_i,\gamma}(k_i))\Sigma_{x \text{ in } X_{t_i,\gamma}(k_i)} f(\theta(k_i, t_i, \alpha_{k_i,t_i}, x)).$$

proof For each (k, t in T(k)), denote by $S(k,t) \subset S$ the subgroup of S generated by the image of $\rho_t(\pi_1(X_t))$. Equivalently, S(k,t) is the subgroup of S generated by the element $A(\mathrm{Frob}_{k,t})$. Denote by $G_{S(k,t)}$ the algebraic group

$$G \subset G_{S(k,t)} \subset G_{arith}$$

which is the inverse image of S(k,t) in G_{arith} under the projection

$$G_{arith} \to G_{arith}/G = S.$$

Denote by $K_{S(k,t)}$ the compact group

$$K \subset K_{S(k,t)} \subset K_{arith}$$

which is the inverse image of S(k,t) under the projection

$$K_{arith} \to K_{arith}/K = S.$$

Thus $K_{S(k,t)}$ is a maximal compact subgroup of $G_{S(k,t)}$. In terms of the cosets $K_{arith,s}$, we have

$$K_{S(k,t)} = \coprod_{s \text{ in } S(k,t)} K_{arith,s}.$$

On X_t, we have the lisse sheaf $\mathcal{F}_t\otimes(\alpha_{k,t})^{\deg}$, whose G_{geom} is G and whose G_{arith} is $G_{S(t,k)}$. We also have the finite etale Γ-torsor Y_t/X_t, and its pushout sheaf $\mathcal{S}_\Gamma$ on X. By assumption 4), $\pi_1(X_t)^{geom}$ maps onto Γ. We have already seen (9.5.7.1) that on each fibre X_t, we have

$$|\#X_{t,\gamma}(k)/\#X_t(k) - \#\gamma/\#\Gamma| \le 2C(X_t\otimes_k\bar{k}, \mathcal{S}_{\Gamma,t})\#\Gamma/(\#k)^{1/2}.$$

By the Uniformity Lemma 9.4.3, the constants $C(X_t\otimes_k\bar{k}, \mathcal{S}_{\Gamma,t})$ are all bounded by some $C(X/T, \mathcal{S}_\Gamma)$, so we get the uniform estimate

$$|\#X_{t,\gamma}(k)/\#X_t(k) - \#\gamma/\#\Gamma| \le 2C(X/T, \mathcal{S}_\Gamma)\#\Gamma/(\#k)^{1/2}.$$

In particular, we have the limit formula

$$\#\Gamma/\#\gamma = \lim_{i \to \infty} \#X_{t_i}(k_i)/\#X_{t_i,\gamma}(k_i).$$

It remains only to show that for any continuous central function $F(k, \gamma)$ on $K_{arith}\times\Gamma$, we have the limit formula

$$\int_{K_{arith,s}\times\Gamma} F(k, g)dkdg = \lim_{i \to \infty} (1/\#X_{t_i}(k_i))\Sigma_{x \text{ in } X_{t_i}(k_i)} F(\theta(k_i, t_i, \alpha_{k_i,t_i}, x), \gamma(k,x)).$$

For then we take $F(k, g) := f(g)I_\gamma(g)$, where I_γ is the characteristic function of the conjugacy class Γ. The above limit formula specializes to

$$(\#\gamma/\#\Gamma)\int_{K_{arith,s}} f(k)dk = \lim_{i \to \infty} (1/\#X_{t_i}(k_i))\Sigma_{x \text{ in } X_{t_i,\gamma}(k_i)} f(\theta(k_i, t_i, \alpha_{k_i,t_i}, x)).$$

One then multiplies this limit formula with the limit formula

$$\#\Gamma/\#\gamma = \lim_{i \to \infty} \#X_{t_i}(k_i)/\#X_{t_i,\gamma}(k_i).$$

above.

How do we show that

$$\int_{K_{arith,s}\times\Gamma} F(k, g)dkdg = \lim_{i \to \infty} (1/\#X_{t_i}(k_i))\Sigma_{x \text{ in } X_{t_i}(k_i)} F(\theta(k_i, t_i, \alpha_{k_i,t_i}, x), \gamma(k,x))$$

for any continuous central function $F(k, g)$ on $K_{arith}\times\Gamma$? It suffices to treat the case when F is the trace of a finite–dimensional representation Λ of $K_{arith}\times\Gamma$.

For each (k, t in T(k)) with $A(Frob_{k,t}) = s$, we apply Deligne's equidistribution theorem [Ka–Sar, RMFEM, 9.7.10] to the sheaf $\mathcal{F}_t\otimes(\alpha_{k,t})^{deg} \oplus \mathcal{S}_{\Gamma,t}$ on X_t. Its G_{geom} is $G\times\Gamma$ and its G_{arith} is $G_{S(k,t)}\times\Gamma$, with compact forms $K\times\Gamma \subset K_{S(k,t)}\times\Gamma$. We restrict the representation Λ to $K_{S(k,t)}\times\Gamma$. For $Card(k) \geq 4A(X/T)^2$, and t in T(k) with $A(Frob_{k,t}) = s$, we have the estimate

$$\left|\int_{K_{arith,s}\times\Gamma} Trace(\Lambda(k,g))dkdg \ - \ (1/\#X_t(k))\Sigma_{x \text{ in } X_t(k)} F(\theta(k, t, \alpha_{k,t}, x), \gamma(k,x))\right|$$
$$\leq 2C(X_t\otimes_k\bar{k}, \mathcal{F}_t\otimes(\alpha_{k,t})^{deg} \oplus \mathcal{S}_{\Gamma,t})dim(\Lambda)/Card(k)^{1/2}.$$

The trivial but key observation here is that on $X_t\otimes_k\bar{k}$, the sheaf $\mathcal{F}_t\otimes(\alpha_{k,t})^{deg}$ is isomorphic to $\mathcal{F}_t$ (because $(\alpha_{k,t})^{deg}$ is geometrically constant). So by the Uniformity Lemma 9.4.3, we have the uniform estimate

$$\left|\int_{K_{arith,s}\times\Gamma} Trace(\Lambda(k,g))dkdg - (1/\#X_t(k))\Sigma_{x \text{ in } X_t(k)} F(\theta(k, t, \alpha_{k,t}, x), \gamma(k,x))\right|$$
$$\leq 2C(X_t\otimes_k\bar{k}, \mathcal{F}_t\oplus\mathcal{S}_{\Gamma,t})dim(\Lambda)/Card(k)^{1/2}$$
$$\leq 2C(X/T, \mathcal{F}\oplus\mathcal{S}_\Gamma)dim(\Lambda)/Card(k)^{1/2}.$$ QED

(9.6.4) Also quite useful is the following special case $\Gamma = \{e\}$ of the above result, which is a slight variant of [Ka–Sar, RMFEM, 9.7.10].

Theorem 9.6.5 Suppose given a prime ℓ, a field embedding $\iota : \bar{\mathbb{Q}}_\ell \to \mathbb{C}$, a connected normal $\mathbb{Z}[1/\ell]$–scheme T of finite type, a smooth T–scheme X/T with geometrically connected fibres of dimension d, and a lisse $\bar{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on X of rank $r \geq 1$. We fix two (not necessarily connected) semisimple $\bar{\mathbb{Q}}_\ell$–algebraic subgroups $G \subset G_{arith}$ of GL(r). We suppose that G is a normal subgroup of G_{arith} of finite index, and that the quotient group G_{arith}/G is a finite cyclic group S.

We fix maximal compact subgroups K in $G(\mathbb{C})$ and K_{arith} in $G_{arith}(\mathbb{C})$, with $K \subset K_{arith}$. We make the following hypothesis:

For every finite field k, and every k–valued point t of T, there exists a constant $\alpha_{k,t}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ such that

1) The lisse sheaf on X_t/k given by $\mathcal{F}_t\otimes(\alpha_{k,t})^{deg}$ is ι–pure of weight zero.

2) The group G_{geom} for $\mathcal{F}_t\otimes(\alpha_{k,t})^{deg}$ is (conjugate in GL(r) to) G.

3) Under the representation ρ_t of $\pi_1(X_t)$ corresponding to $\mathcal{F}_t\otimes(\alpha_{k,t})^{deg}$, the entire group $\pi_1(X_t)$ lands in G_{arith}, i.e., we have $\rho_t(\pi_1(X_t)) \subset G_{arith}$.

4) There exists a surjective homomorphism

$$a : \pi_1(T) \to S$$

with the following property. For each finite field k, each k–valued point t in T(k), and each k–valued point x in $X_t(k)$, the image in S of $\rho_t(\mathrm{Frob}_{k,x})$ is $A(\mathrm{Frob}_{k,t})$.

With these hypotheses, fix an element s in S. For each finite field k, each k–valued point t of T with $A(\mathrm{Frob}_{k,t}) = s$, and each k–valued point x of X_t, denote by $\theta(k, t, \alpha_{k,t}, x)$ the Frobenius conjugacy class in $K_{arith,s}$ attached to the point x and the lisse sheaf $\mathcal{F}_t\otimes(\alpha_{k,t})^{deg}$ on X_t. Fix a continuous $\mathbb{C}$–valued central function f on K_{arith}. Fix any sequence of data (k_i, t_i in $T(k_i)$) in which the k_i are finite fields with

$$\mathrm{Card}(k_i) > 4A(X/T)^2$$

whose cardinalities form a strictly increasing sequence, and in which, for each i, $A(\mathrm{Frob}_{k_i,t_i}) = s$.

We have the limit formula

$$\int_{K_{arith,s}} f(k)dk = \lim_{i \to \infty} (1/\#X_{t_i}(k_i))\Sigma_{x \text{ in } X_{t_i}(k_i)}\, f(\theta(k_i, t_i, \alpha_{k_i,t_i}, x)).$$

More precisely, for Λ any finite–dimensional representation of K_{arith}, any finite field k with $\mathrm{Card}(k) \geq 4A(X\otimes_k\overline{k})^2$, and any t in T(k) with $A(\mathrm{Frob}_{k,t}) = s$, we have the estimate

$$|\int_{K_{arith,s}} \mathrm{Trace}(\Lambda(k))dk - (1/\#X_t(k))\Sigma_{x \text{ in } X_t(k)}\, \mathrm{Trace}(\Lambda(\theta(k, t, \alpha_{k,t}, x)))|$$
$$\leq 2C(X\otimes_k\overline{k}, \mathcal{F})\dim(\Lambda)/\mathrm{Card}(k)^{1/2}.$$

proof Take Γ to be the trivial group in Theorem 9.6.3. QED

9.7 Applications of Goursat's Lemma

(9.7.1) We now explore some conditions which guarantee that the group $G_{geom,\mathcal{F}\times\Gamma}$ is equal to the product $G_{geom,\mathcal{F}}\times \Gamma$. The key point is that $G_{geom,\mathcal{F}\times\Gamma}$ is a Zariski–closed subgroup of

$G_{geom,\mathcal{F}} \times \Gamma$ which maps onto both factors.

Lemma 9.7.2 (Goursat) Let $G/\mathbb{C}$ be an algebraic group of finite type over an algebraically closed field of characteristic zero. Let Γ be a finite group (viewed as algebraic group over $\mathbb{C}$ by means of some faithful linear representation). Let H be a Zariski–closed subgroup of $G\times\Gamma$ which maps onto each factor. Then there exists a closed normal subgroup G_1 of G with $G^0 \subset G_1$, and a normal subgroup $\Gamma_1 \subset \Gamma$, such that H is the inverse image in $G\times\Gamma$ of the graph of an isomorphism between G/G_1 and Γ/Γ_1.

proof Since H maps onto G, $\dim(H) \geq \dim(G)$. But $H \subset G\times\Gamma$ with Γ finite, so $\dim(H) \leq \dim(G\times\Gamma) = \dim(G)$. Therefore $\dim(H) = \dim(G\times\Gamma)$. As H is a closed subgroup of $G\times\Gamma$, the identity component H^0 of H must be the identity component $(G\times\Gamma)^0 = G^0\times\{e\}$ of $G\times\Gamma$. Therefore H contains $G^0\times\{e\}$. So H is the inverse image in $G\times\Gamma$ of some subgroup $\overline{H}$ of the finite group $(G/G^0)\times\Gamma$ which maps onto both factors of $(G/G^0)\times\Gamma$. This reduces us to treating universally the case when the group G is finite, in which case this is the classical Goursat Lemma, cf. [Lang, Algebra, ex. 5 on page 75]. QED

Corollary 9.7.3 Hypotheses as in 9.7.2, if G is connected, then H is $G\times\Gamma$.

Corollary 9.7.4 Hypotheses as in 9.7.2, suppose Γ is the symmetric group S_d with $d \geq 5$. If G/G^0 has no quotient group of order two, then H is $G\times\Gamma$.

proof Either H is $G\times\Gamma$ or it is the inverse image of the graph of an isomorphism between a nontrivial quotient of G/G^0 and a nontrivial quotient of S_d. The only nontrivial quotients of S_d are S_d itself and $S_d/A_d \cong \{\pm 1\}$, both of which admit quotients of order 2. Since G/G^0 admits no such quotient, we must have $H = G\times\Gamma$ by the paucity of choice. QED

9.8 Interlude: detailed discussion of the $O(N)\times S_d$ case

(9.8.1) This last corollary, 9.7.4, is of no use if G is the orthogonal group $O(N)$, and Γ is S_d with $d \geq 5$.

Theorem 9.8.2 Let k be a field, X/k a smooth, geometrically connected k–scheme, ξ a geometric point of X, ℓ a prime invertible in k, and $\mathcal{F}$ a lisse, orthogonally self–dual $\overline{\mathbb{Q}}_\ell$–sheaf on X of some rank N, corresponding to a representation

$$\Theta_{\mathcal{F}}: \pi_1(X, \xi) \to O(N),$$

whose G_{geom} is the full orthogonal group $O(N)$. Let

$$\rho : \pi_1{}^{\mathrm{geom}}(X, \xi) \twoheadrightarrow S_d$$

be a surjective homomorphism onto the symmetric group S_d for some $d \geq 5$. Let us denote by

$$\mathrm{sgn}(\rho) : \pi_1{}^{\mathrm{geom}}(X, \xi) \twoheadrightarrow S_d/A_d = \{\pm 1\}$$

the $\{\pm 1\}$–valued character of $\pi_1{}^{\mathrm{geom}}(X, \xi)$ obtained by composing ρ with the sign character of S_d, and by

$$\mathcal{L}_{\mathrm{sgn}(\rho)}$$

the corresponding lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaf on $X \otimes_k \overline{k}$. Then we have the following possibilities for $G_{\mathrm{geom}, \mathcal{F} \times \rho}$.

1) Suppose that the lisse rank one $\overline{\mathbb{Q}}_\ell$–sheaves $\mathcal{L}_{\mathrm{sgn}(\rho)}$ and $\det(\mathcal{F})$ are isomorphic on $X \otimes_k \overline{k}$, i.e., suppose that the two $\{\pm 1\}$–valued characters of $\pi_1{}^{\mathrm{geom}}(X, \xi)$ given by $\mathrm{sgn}(\rho)$ and by $\det(\Theta_{\mathcal{F}})$ are equal. Then $G_{\mathrm{geom}, \mathcal{F} \times \rho}$ is the subgroup of $O(N) \times S_d$ of all elements (A, σ) with $\det(A) = \mathrm{sgn}(\sigma)$.

2) Suppose that $\mathcal{L}_{\mathrm{sgn}(\rho)}$ and $\det(\mathcal{F})$ are not geometrically isomorphic, i.e., suppose that $\mathrm{sgn}(\rho) \neq \det(\Theta_{\mathcal{F}})$ as characters of $\pi_1{}^{\mathrm{geom}}(X, \xi)$. Then $G_{\mathrm{geom}, \mathcal{F} \times \rho}$ is the entire product $O(N) \times S_d$.

proof Since the only nontrivial quotient of $O(N)/O(N)^0 = \{\pm 1\}$ is $\{\pm 1\}$, and S_d has unique quotient $\{\pm 1\}$ by the sign character, either $G_{\mathrm{geom}, \mathcal{F} \times \rho}$ is the entire product $O(N) \times S_d$, or it is the subgroup of $O(N) \times S_d$ consisting of all elements (A, σ) with $\det(A) = \mathrm{sgn}(\sigma)$.

In the latter case, the characters $(A, \sigma) \mapsto \det(A)$ and $(A, \sigma) \mapsto \mathrm{sgn}(\sigma)$ coincide on $G_{\mathrm{geom}, \mathcal{F} \times \rho}$. In particular these characters coincide on elements $(\Theta_{\mathcal{F}}(\gamma), \rho(\gamma))$ with γ in $\pi_1{}^{\mathrm{geom}}(X, \xi)$. This means exactly that $\det(\Theta_{\mathcal{F}}) = \mathrm{sgn}(\rho)$ on $\pi_1{}^{\mathrm{geom}}(X, \xi)$.

In the former case, the two characters $\det(\Theta_{\mathcal{F}})$ and $\mathrm{sgn}(\rho)$ on $\pi_1{}^{\mathrm{geom}}(X, \xi)$ must be distinct, otherwise by the Zariski density of $(\Theta_{\mathcal{F}} \times \rho)(\pi_1{}^{\mathrm{geom}}(X, \xi))$ the two characters $(A, \sigma) \mapsto \det(A)$ and $(A, \sigma) \mapsto \mathrm{sgn}(\sigma)$ would coincide on $G_{\mathrm{geom}, \mathcal{F} \times \rho} = O(N) \times S_d$, which they do not. QED

9.9 Application to twist sheaves

Theorem 9.9.1 Let k be an algebraically closed field in which 2 is invertible. Fix a prime number ℓ which is invertible in k. Denote by χ_2 the unique character of order 2 of the tame fundamental group of $\mathbb{G}_m/k$. Let C/k be a proper smooth connected curve of genus g. Fix an irreducible middle extension $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on C, which is symplectically self–dual. Let $D = \sum a_i P_i$ be an effective divisor of degree d on C. Suppose that

1) $\quad d \geq 4g+4,$
and
$$2g - 2 + d > \mathrm{Max}(2\#\mathrm{Sing}(\mathcal{F}), 72\mathrm{rank}(\mathcal{F})).$$

2) Either $\mathcal{F}$ is everywhere tame, or $\mathcal{F}$ is tame at all points of D and the characteristic p is either zero or $p \geq \mathrm{rank}(\mathcal{F}) + 2$.
3) There exists a finite singularity β of $\mathcal{F}$, i.e., a point β in $\mathrm{Sing}(\mathcal{F})\cap(C-D)$, such that the following two conditions hold:

3a) $\mathcal{F}$ is tame at β,

3b) $\mathcal{F}(\beta)/\mathcal{F}(\beta)^{I(\beta)}$ has odd dimension.

Consider the lisse sheaf $\mathcal{G}$ on $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ given by
$$f \mapsto H^1(C, j_*(\mathcal{F}\otimes\mathcal{L}_{\chi(f)}),$$
whose G_{geom} is, by Theorem 5.5.1, the full orthogonal group O. The lisse rank one sheaf $\det(\mathcal{G})$ on $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ is not the restriction to $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ of a lisse sheaf on the larger space $Fct(C, d, D, \emptyset)$.

proof Pick f_1 and f_2 as in the proof of 5.4.9. Consider the pullback
$$\mathcal{H} := [t \mapsto f_1(t - f_2)]^*\mathcal{G}$$
to $\mathbb{A}^1 - \mathrm{CritVal}(f_2, \mathcal{F}\otimes\mathcal{L}_{\chi(f_1)})$. At the point $t = f_2(\beta)$, $\det(\mathcal{H})$ has nontrivial local monodromy (the character of order two), cf 5.4.11.

On the other hand, the function $f_1(f_2(\beta) - f_2)$ on C has d distinct zeroes, all disjoint from D, i.e., the function $f_1(f_2(\beta) - f_2)$ lies in $\mathrm{Fct}(C, d, D, \emptyset)$. To see this, recall that f_1 was chosen to lie in $\mathrm{Fct}(C, \deg(D_1), D_1, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}})$, so f_1 has d_1 distinct zeroes, all disjoint from D. Then f_2 was required to lie in $\mathrm{Fct}(C, \deg(D_2), D_2, \mathrm{Sing}(\mathcal{F})\cup D^{\mathrm{red}}\cup f_1{}^{-1}(0))$ and to lie in the open set U of Theorem 2.2.6 with respect to the set

$S := f_1{}^{-1}(0)\cup(\mathrm{Sing}(\mathcal{F})\cap(C-D_2)).$

The point β lies in S, so $f_2 - f_2(\beta)$ has d_2 distinct zeroes, all disjoint from D. Also, f_2 is injective on the set $f_1{}^{-1}(0)\cup(\mathrm{Sing}(\mathcal{F})\cap(C-D_2))$, so it is injective on the subset $f_1{}^{-1}(0)\cup\{\beta\}$. Therefore $f_2(\beta) - f_2$ is nonzero at every zero of f_1. Thus, $f_1(f_2(\beta) - f_2)$ has d distinct zeroes, all disjoint from D. In other words, $f_1(f_2(\beta) - f_2)$ lies in $\mathrm{Fct}(C, d, D, \emptyset)$.

Now suppose there exists a lisse sheaf $\mathcal{L}$ on $Fct(C, d, D, \emptyset)$ whose restriction to $Fct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\mathrm{finite}})$ is $\det(\mathcal{G})$. Then the pullback $[t \mapsto f_1(t - f_2)]^*\mathcal{L}$ is lisse at $t=f_2(\beta)$, precisely because the function $f_1(f_2(\beta) - f_2)$ lies in Fct(C, d, D, Ø). But this same pullback is $\det(\mathcal{H})$, which is not lisse at $f_2(\beta)$, contradiction. QED

Corollary 9.9.2 Notations and hypotheses as in Theorem 9.9.1 above, denote by

$$\rho_{\text{split}} : \pi_1(\mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}}), \xi) \to S_d$$

the homomorphism attached to the finite etale S_d–torsor

$$\mathrm{Split}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}}) \to \mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}}).$$

Then $G_{\text{geom},\mathcal{G}\times\rho_{\text{split}}}$ is the product group $O\times S_d$.

proof In view of Theorem 9.8.2, we need only show that $\det(\mathcal{G})$ is not isomorphic to $\mathcal{L}_{\mathrm{sgn}(\rho_{\text{split}})}$ as lisse sheaf on the space

$$\mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}}).$$

But $\mathcal{L}_{\mathrm{sgn}(\rho_{\text{split}})}$ is the restriction to $\mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}})$ of a lisse sheaf on $\mathcal{F}ct(C, d, D, Ø)$, since the finite etale S_d–torsor

$$\mathrm{Split}(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}}) \to \mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}})$$

is the restriction to $\mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}})$ of the finite etale S_d–torsor

$$\mathrm{Split}(C, d, D, Ø) \to \mathcal{F}ct(C, d, D, Ø).$$

In view of the above theorem, $\det(\mathcal{G})$ is not such a restriction, hence cannot be isomorphic to $\mathcal{L}_{\mathrm{sgn}(\rho_{\text{split}})}$ as lisse sheaf on $\mathcal{F}ct(C, d, D, \mathrm{Sing}(\mathcal{F})_{\text{finite}})$. QED

9.10 Equidistribution theorems for twists by primes, over finite fields

(9.10.1) In this section, we put ourselves in the situation of 7.0, and give ourselves data $(C/k, D, \ell, r, \mathcal{F}, \chi, \iota, w)$. We suppose that after extension of scalars from k to $\overline{k}$, our data $(C/k, D, \ell, r, \mathcal{F}, \chi)$ satisfies all the hypotheses of Theorem 5.5.1 or of Theorem 5.6.1.

Theorem 9.10.2 Hypotheses as in 9.10.1 above, suppose that G_{geom} for $\mathcal{G}$ is the group $SL_\nu(N)$ for some **odd** integer ν. Choose β such that $\mathcal{G}\otimes\beta^{\deg}$ is ι–pure of weight zero, and all its Frobenii land in G_{geom}.

1) Fix a conjugacy class σ in the symmetric group S_d. As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\sigma-\text{split}}(E)}$ become equidistributed for Haar measure in the space $U_\nu(N)^\#$ of conjugacy classes in $U_\nu(N)$.

2) As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\text{prime}}(E)}$ become equidistributed for Haar measure in the space $U_\nu(N)^\#$ of

conjugacy classes in $U_\nu(N)$.

proof Assertion 1) results from 9.7.4 and 9.5.11. Assertion 2) is the special case of 1) in which we take for σ the class of a d–cycle. QED

Theorem 9.10.3 Hypotheses as in 9.10.1 above, suppose that $\mathcal{G}((w+1)/2)$ is symplectically self–dual on X, and suppose that G_{geom} for $\mathcal{G}$ is the group Sp(N).
1) Fix a conjugacy class σ in the symmetric group S_d. As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\sigma-\mathrm{split}}(E)}$ become equidistributed for Haar measure in the space $USp(N)^\#$ of conjugacy classes in USp(N).
2) As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\mathrm{prime}}(E)}$ become equidistributed for Haar measure in the space $USp(N)^\#$ of conjugacy classes in USp(N).

proof Assertion 1) results from 9.7.3 and 9.5.11. Assertion 2) is the special case of 1) in which we take for σ the class of a d–cycle. QED

Theorem 9.10.4 Hypotheses as in 9.10.1 above, suppose that $\mathcal{G}((w+1)/2)$ is orthogonally self–dual on X. Suppose that $G_{geom} = G_{arith} = SO(N)$ for $\mathcal{G}((w+1)/2)$.
1) Fix a conjugacy class σ in the symmetric group S_d. As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\sigma-\mathrm{split}}(E)}$ become equidistributed for Haar measure in the space $SO(N)^\#$ of conjugacy classes in SO(N).
2) As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\mathrm{prime}}(E)}$ become equidistributed for Haar measure in the space $SO(N)^\#$ of conjugacy classes in SO(N).

proof Assertion 1) results from 9.7.3 and 9.5.11. Assertion 2) is the special case of 1) in which we take for σ the class of a d–cycle. QED

Theorem 9.10.5 Hypotheses as in 9.10.1 above, suppose that $\mathcal{G}((w+1)/2)$ is orthogonally self–dual on X, and suppose that G_{geom} for $\mathcal{G}$ is the group O(N). Suppose further that $\det(\mathcal{G})$ on $X\otimes_k\overline{k}$ is not the restriction to $X\otimes_k\overline{k}$ of a lisse sheaf on $\mathcal{F}ct(C, d, D, \emptyset)\otimes_k\overline{k}$, cf. 9.9.1 for examples.
1) Fix a conjugacy class σ in the symmetric group S_d. As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\sigma-\mathrm{split}}(E)}$ become equidistributed for Haar measure in the space $O(N)^\#$ of conjugacy classes in O(N).

2) As E runs over finite extensions of k with $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{prime}(E)}$ become equidistributed for Haar measure in the space $O(N)^{\#}$ of conjugacy classes in O(N).

proof Assertion 1) results from 9.8.2, 9.9.2, and 9.5.11. Assertion 2) is the special case of 1) in which we take for σ the class of a d−cycle. QED

(9.10.6) Let us spell this out in terms of the decomposition

$$O(N, \mathbb{R})^{\#} = O_{sign\ +}(N, \mathbb{R})^{\#} \amalg O_{sign\ -}(N, \mathbb{R})^{\#}.$$

Corollary 9.10.7 Hypotheses as in 9.10.5, we have:

1) Fix a conjugacy class σ in S_d, and a sign $\varepsilon = \pm 1$. For each finite extension E/k, denote by $X_{sign\ \varepsilon}(E)$ the subset of X(E) consisting of those points f in X(E) such that

$$\det(-\mathrm{Frob}_{E,f} \mid \mathcal{G}((w+1)/2)) = \varepsilon.$$

Denote by $X_{sign\ \varepsilon,\ \sigma-split}(E)$ the subset of X(E) given by

$$X_{sign\ \varepsilon,\ \sigma-split}(E) := X_{sign\ \varepsilon}(E) \cap X_{\sigma-split}(E).$$

As $\#E \to \infty$,

$$\#X_{sign\ \varepsilon}(E)/\#X(E) \to (1/2)\times(\#\sigma/d!),$$

and the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{sign\ \varepsilon,\ \sigma-split}(E)}$ become equidistributed for Haar measure of total mass one on the space $O_{sign\ \varepsilon}(N, \mathbb{R})^{\#}$.

2) Fix a sign $\varepsilon = \pm 1$. As $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{sign\ \varepsilon,\ prime}(E)}$ become equidistributed for Haar measure of total mass one on the space $O_{sign\ \varepsilon}(N, \mathbb{R})^{\#}$.

proof Assertion 1) is obtained by applying the equidistribution statement 1) of Theorem 9.10.5 to the integration of continuous central functions on $O(N, \mathbb{R})$ which are supported in $O_{sign\ \varepsilon}(N, \mathbb{R})$. Assertion 2) is the special case of 1) where we take for σ the class of a d−cycle. QED

Theorem 9.10.8 Hypotheses as in 9.10.1, suppose that $\mathcal{G}((w+1)/2)$ is orthogonally self−dual on X. Suppose that $\mathcal{G}((w+1)/2)$ has $G_{geom} = SO(N)$ and $G_{arith} = O(N)$. Then we have:

1) The rank N of $\mathcal{G}$ is even.

2) Fix a conjugacy class σ in the symmetric group S_d, and a sign $\varepsilon = \pm 1$. As E runs over finite extensions of k with $(-1)^{\deg(E/k)} = \varepsilon$ and $\#E \to \infty$, the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{\sigma-split}(E)}$ become equidistributed for Haar measure in the space $O_{sign\ \varepsilon}(N, \mathbb{R})^{\#}$.

3) Fix a sign $\varepsilon = \pm 1$. As E runs over finite extensions of k with $\#E \to \infty$ and $(-1)^{\deg(E/k)} = \varepsilon$,

the conjugacy classes $\{\theta(E, f)\}_{f \text{ in } X_{prime}(E)}$ become equidistributed for Haar measure in the space $O_{sign\ \varepsilon}(N, \mathbb{R})^\#$.

proof Assertion 1) results from 5.5.2, part 3). Assertion 2) results from 9.7.3 and 9.5.11. Assertion 3) is the special case of 2) where we take for σ the class of a d−cycle. QED

9.11 Average analytic ranks of twists by primes over finite fields

(9.11.1) We first give the result in the case when G_{geom} is the full orthogonal group.

Theorem 9.11.2 Hypotheses and notations as in Theorem 9.10.5 and Corollary 9.10.7 above, fix a conjugacy class σ in the symmetric group S_d. If we take the limit over finite extensions E/k large enough that the sets $X_{\sigma-split}(E)$ and $X_{sign\ \varepsilon,\ \sigma-split}(E)$ are all nonempty, we get the following tables of limit formulas. In these tables, the number in the third column is the limit, as $\#E \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{\sigma-split}(E)$	$rank_{an}(\mathcal{G}, E, f)$	1/2,
$X_{\sigma-split}(E)$	$rank_{quad,\ an}(\mathcal{G}, E, f)$	1,
$X_{\sigma-split}(E)$	$rank_{geom,\ an}(\mathcal{G}, E, f)$	1.

More precisely, when we break up $X_{\sigma-split}(E)$ according to the sign ε in the functional equation, we have the following tables of limit values (same format as above).

if N is even

$X_{sign\ -,\ \sigma-split}(E)$	$rank_{an}(\mathcal{G}, E, f)$	1,
$X_{sign\ +,\ \sigma-split}(E)$	$rank_{an}(\mathcal{G}, E, f)$	0,

$X_{sign\ -,\sigma-split}(E)$	$rank_{quad,\ an}(\mathcal{G}, E, f)$	2,
$X_{sign\ +,\sigma-split}(E)$	$rank_{quad,\ an}(\mathcal{G}, E, f)$	0,

$X_{sign\ -,\sigma-split}(E)$	$rank_{geom,\ an}(\mathcal{G}, E, f)$	2,
$X_{sign\ +,\sigma-split}(E)$	$rank_{geom,\ an}(\mathcal{G}, E, f)$	0.

if N is odd

$X_{\text{sign } -,\ \sigma-\text{split}}(E)$	$\text{rank}_{\text{an}}(\mathcal{G}, E, f)$	1,
$X_{\text{sign } +,\ \sigma-\text{split}}(E)$	$\text{rank}_{\text{an}}(\mathcal{G}, E, f)$	0,
$X_{\text{sign } -,\sigma-\text{split}}(E)$	$\text{rank}_{\text{quad, an}}(\mathcal{G}, E, f)$	1,
$X_{\text{sign } +,\sigma-\text{split}}(E)$	$\text{rank}_{\text{quad, an}}(\mathcal{G}, E, f)$	1,
$X_{\text{sign } -,\sigma-\text{split}}(E)$	$\text{rank}_{\text{geom, an}}(\mathcal{G}, E, f)$	1,
$X_{\text{sign } +,\sigma-\text{split}}(E)$	$\text{rank}_{\text{geom, an}}(\mathcal{G}, E, f)$	1.

If we take σ to be the conjugacy class of a d−cycle in S_d, then the set $X_{\sigma-\text{split}}(E)$ becomes $X_{\text{prime}}(E)$, and $X_{\text{sign } \varepsilon,\ \sigma-\text{split}}(E)$ becomes $X_{\text{sign } \varepsilon,\ \text{prime}}(E)$.

proof Combine the equidistribution statements of Theorem 9.10.5 and Corollary 9.10.7 with the proof of Theorem 8.3.3. QED

(9.11.3) We now give the analogous result in the remaining cases.

Theorem 9.11.4 Hypotheses and notations as in Theorem 9.10.4, fix a conjugacy class σ in the symmetric group S_d. For every finite extension E/k, $X_{\text{sign } -}(E)$ is empty. If we take the limit over finite extensions E/k large enough that $X_{\sigma-\text{split}}(E) = X_{\text{sign } +,\sigma-\text{split}}(E)$ is nonempty, we get the following tables of limit formulas. In these tables, the number in the third column is the limit, as $\#E \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{\sigma-\text{split}}(E)$	$\text{rank}_{\text{an}}(\mathcal{G}, E, f)$	0,
$X_{\sigma-\text{split}}(E)$	$\text{rank}_{\text{quad, an}}(\mathcal{G}, E, f)$	0,
$X_{\sigma-\text{split}}(E)$	$\text{rank}_{\text{geom, an}}(\mathcal{G}, E, f)$	0.

proof Combine the equidistribution statement of Theorem 9.10.4 with the proof of 8.3.6. QED

Theorem 9.11.5 Hypotheses and notations as in Theorem 9.10.8, fix a conjugacy class σ in the symmetric group S_d, and a sign $\varepsilon = \pm 1$. For every finite extension E/k with $(-1)^{\deg(E/k)} = \varepsilon$, we have $X_{\text{sign } \varepsilon}(E) = X(E)$, and $X_{\text{sign } -\varepsilon}(E)$ is empty. If we take the limit all finite extensions E/k with $(-1)^{\deg(E/k)} = \varepsilon$ and large enough that the sets $X_{\sigma-\text{split}}(E) = X_{\text{sign } \varepsilon,\sigma-\text{split}}(E)$ are all nonempty, we get the following tables of limit formulas. In these tables, the number in the third column is the limit, as $\#E \to \infty$ over fields E/k with $(-1)^{\deg(E/k)} = \varepsilon$, of the average value of the quantity in the second column over all f's in the set named in the first column.

$\varepsilon = -1$

$X_{\text{sign } -,\, \sigma-\text{split}}(E)$	$\mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E, f)$	1,
$X_{\text{sign } -,\sigma-\text{split}}(E)$	$\mathrm{rank}_{\mathrm{quad,\, an}}(\mathcal{G}, E, f)$	2,
$X_{\text{sign } -,\sigma-\text{split}}(E)$	$\mathrm{rank}_{\mathrm{geom,\, an}}(\mathcal{G}, E, f)$	2,

$\varepsilon = +1$

$X_{\text{sign } +,\, \sigma-\text{split}}(E)$	$\mathrm{rank}_{\mathrm{an}}(\mathcal{G}, E, f)$	0,
$X_{\text{sign } +,\sigma-\text{split}}(E)$	$\mathrm{rank}_{\mathrm{quad,\, an}}(\mathcal{G}, E, f)$	0,
$X_{\text{sign } +,\sigma-\text{split}}(E)$	$\mathrm{rank}_{\mathrm{geom,\, an}}(\mathcal{G}, E, f)$	0.

proof Combine the equidistribution statement of Theorem 9.10.8 with the proof of 8.3.8. QED

Chapter 10: Horizontal Results

10.0 The basic horizontal setup

(10.0.1) We fix a prime number ℓ, an integer $n \geq 2$, and a character

$$\chi : \mu_n(\mathbb{Z}[1/\ell n, \zeta_n]) \to (\overline{\mathbb{Q}}_\ell)^\times$$

of order n. We fix a **nonempty** connected normal $\mathbb{Z}[1/\ell n, \zeta_n]$–scheme T of finite type. We fix a proper, smooth, geometrically connected curve C/T of genus g. We suppose given an effective Cartier divisor S in C which is finite etale over T of degree $s \geq 0$ (with the convention that S is empty if $e = 0$). We suppose given a lisse $\overline{\mathbb{Q}}_\ell$–sheaf $\mathcal{F}$ on $C - S$ of rank $r \geq 1$. If n is 4 or 6, we suppose that $r \leq 2$. We suppose given an integer w, and a field embedding $\iota : \overline{\mathbb{Q}}_\ell \to \mathbb{C}$, such that $\mathcal{F}$ is ι–pure of weight w. We suppose that for each geometric point t in T, the following three conditions are satisfied:

1) the sheaf $\mathcal{F}_t := \mathcal{F}|(C-S)_t$ on $(C-S)_t$ is geometrically irreducible,

2) denoting by $j_t : (C-S)_t \to C_t$ the inclusion, the irreducible middle extension $(j_t)_*\mathcal{F}_t$ on C_t is not lisse at any point of Sing_t, i.e.,

$$S_t = \mathrm{Sing}((j_t)_*\mathcal{F}_t),$$

3) either $\mathcal{F}_t$ is tame at each point of S_t, or $(r+1)!$ is invertible in the residue field $\kappa(t)$ at t.

(10.0.2) We further suppose that for variable geometric points t in T, the Euler characteristic

$$\chi_c((C-S)_t, \mathcal{F}_t)$$

is a constant function of t. Recall [Ka–SE, 4.7.1] that if the generic point of T has characteristic zero, then each $\mathcal{F}_t$ is automatically everywhere tame, and the Euler characteristic $\chi_c((C-S)_t, \mathcal{F}_t)$ is constant, given by

$$\chi_c((C-S)_t, \mathcal{F}_t) = (2 - 2g - s)r.$$

(10.0.3) Given an effective Cartier divisor D in C, finite and flat over T of degree d, we say that D is adapted to the data (C/T, S, $\mathcal{F}$) if, etale locally on T, we have the following situation.

1) There are pairwise disjoint sections P_i of C/T such that D is $\sum a_i P_i$ for some strictly positive integers a_i with $\sum a_i = d$.

2) There are pairwise disjoint sections Q_j of C/T such that S is $\sum Q_j$, and, for each pair (i, j), either $P_i = Q_j$ or P_i is disjoint from Q_j.

3) $\mathcal{F}$ is tamely ramified along each section P_i which lies in S. [Notice that $\mathcal{F}$ is lisse near any P_i which does not lie in S.]

(10.0.4) If all these conditions are satisfied, then for variable geometric points t in T, the Euler characteristic

$$\chi_c((C-S-D)_t, \mathcal{F}_t)$$

is a constant function of t. If in addition $d \geq 2g+1$, then Proposition 6.2.10 applies to the data $(C/T, D, S - S\cap D, \ell, r, \mathcal{F}|(C - D - S), \chi)$, and so we may form the lisse sheaf $\mathcal{G}$ on the smooth T−scheme

$$X := \mathit{Fct}(C, d, D, S - S\cap D)/T.$$

We denote this sheaf

$$\mathcal{G} := \mathrm{Twist}_{\chi,C/T,D}(\mathcal{F}).$$

Because $\mathcal{F}$ is ι−pure of weight w, $\mathcal{G}$ is ι−pure of weight w+1.

(10.0.5) We suppose given a sequence of effective Cartier divisors D_ν in C, D_ν finite and flat over T of degree $d_\nu \geq 1$, with the degrees d_ν strictly increasing

$$d_1 < d_2 < d_3 \ldots < d_\nu < d_{\nu+1} < \ldots$$

such that each D_ν is adapted to the data $(C/T, S, \mathcal{F})$. Suppose that each d_ν is large enough that the following inequalities hold:

$$d_\nu \geq 12g + 7,$$
$$d_\nu \geq \mathrm{Max}(6g+9, 6s + 11),$$
$$2g - 2 + d_\nu > \mathrm{Max}(2s, 72r).$$

(10.0.6) For each ν, Proposition 6.2.10 applies, and we form the lisse sheaf

$$\mathcal{G}_\nu := \mathrm{Twist}_{\chi,C/T,D_\nu}(\mathcal{F})$$

of rank

$$N_\nu \geq r(2g - 2 + d_\nu)$$

on the smooth T−scheme

$$X_\nu := \mathit{Fct}(C, d, D_\nu, S - S\cap D_\nu)/T.$$

The sheaf $\mathcal{G}_\nu$ is ι−pure of weight w+1.

(10.0.7) For each geometric point t of T of residue characteristic not 2 [resp. 2], and each ν, the data $(C_t, D_t, \ell, r, (j_t)_*\mathcal{F}_t, \chi)$ satisfies all the hypotheses of Theorem 5.5.1 [resp. Theorem 5.6.1]. So for the sheaf $\mathcal{G}_{\nu,t}$ on $X_{\nu,t} := \mathit{Fct}(C_t, d, D_{\nu,t}, S_t - S_t\cap D_{\nu,t})$, its group G_{geom} either contains $SL(N_\nu)$, or is equal to one of $SO(N_\nu)$ or $O(N_\nu)$ or, if N_ν is even, $Sp(N_\nu)$.

Autoduality Lemma 10.0.8 Given data $(C/T, S, \mathcal{F})$ as in 10.0.1 above, suppose in addition that for all geometric points t of T, $\mathcal{F}_t$ is everywhere tame (a condition which holds automatically if the generic point of T has characteristic zero). Then the following conditions are equivalent.

1) For every geometric point t of T, the irreducible lisse sheaf $\mathcal{F}_t$ on $C_t - S_t$ is self−dual [resp. orthogonally self−dual, resp. symplectically self−dual].

2) There exists a geometric point t of T such that the irreducible lisse sheaf $\mathcal{F}_t$ on $C_t - S_t$ is self−dual [resp. orthogonally self−dual, resp. symplectically self−dual].

proof We first prove the equivalence of 1) and 2) for self−duality alone. Fix a geometric point t in

T. Since $\mathcal{F}_t$ is irreducible on $C_t - S_t$, it is self–dual if and only if there exists a nonzero sheaf map from $\mathcal{F}_t$ to its dual $(\mathcal{F}_t)^\vee$ (for by the irreducibility, any such nonzero map must be an isomorphism), or, equivalently, if and only if there exists a nonzero sheaf map from $(\mathcal{F}_t)^\vee$ to $\mathcal{F}_t$. Thus $\mathcal{F}_t$ is self–dual if and only if the cohomology group $H^0(C_t - S_t, (\mathcal{F}_t \otimes \mathcal{F}_t)^\vee)$ is nonzero, or, equivalently (Poincaré duality), if and only if the compactly supported cohomology group $H_c{}^2(C_t - S_t, \mathcal{F}_t \otimes \mathcal{F}_t)$ is nonzero.

Denote by $\pi : C - S \to T$ the structural morphism. By proper base change, $H_c{}^2(C_t - S_t, \mathcal{F}_t \otimes \mathcal{F}_t)$ is the stalk at t of the sheaf $R^2\pi_!(\mathcal{F} \otimes \mathcal{F})$. By Deligne's semicontinuity theorem [Lau–SC], the tameness of each $\mathcal{F}_t$, and hence of each $\mathcal{F}_t \otimes \mathcal{F}_t$, on $C_t - S_t$, guarantees that all the higher direct images $R^i\pi_!(\mathcal{F} \otimes \mathcal{F})$ are lisse sheaves on T. As T is connected, the lisse sheaf $R^2\pi_!(\mathcal{F} \otimes \mathcal{F})$ on T vanishes if and only if its stalk at a single point vanishes.

Suppose now that $\mathcal{F}_t$ is self–dual. It is orthogonally self–dual if and only if $H_c{}^2(C_t - S_t, \mathrm{Sym}^2(\mathcal{F}_t))$ is nonzero, and it is symplectically self–dual if and only if $H_c{}^2(C_t - S_t, \Lambda^2(\mathcal{F}_t))$ is nonzero. Once again, both $\mathrm{Sym}^2(\mathcal{F}_t)$ and $\Lambda^2(\mathcal{F}_t)$ are tame, so both $R^2\pi_!(\mathrm{Sym}^2(\mathcal{F}))$ and $R^2\pi_!(\Lambda^2(\mathcal{F}))$ are lisse on T, and we conclude as above. QED

Theorem 10.0.9 Hypotheses and notations as in 10.0.1–5, pick a finite extension E_λ of $\mathbb{Q}_\ell$ which contains the n'th roots of unity (n := the order of χ), and large enough that $\mathcal{F}$ has an E_λ–form. [Thus each $\mathcal{G}_\nu$ has an E_λ–form.] Denote by $\mu(E_\lambda)$ the number of roots of unity in E_λ. Then we have the following results.

1) (the SL case) Suppose that either $n \geq 3$, or that for every geometric point t of T, $\mathcal{F}_t$ is not self–dual. Then for each ν, and for each geometric point t in T, there exists a divisor $m_{\nu,t}$ of $\mu(E_\lambda)$ such that the group G_{geom} for $\mathcal{G}_{\nu,t}$ is $GL_{m_{\nu,t}}(N_\nu)$. Moreover, for each ν there exists a dense open set U_ν of T on which the function $t \mapsto m_{\nu,t}$ is constant, say with value m_ν. Every $m_{\nu,t}$ divides the generic value m_ν.

2) (the Sp case) Suppose that χ has order 2, and that for every geometric point t of T, $\mathcal{F}_t$ is orthogonally self–dual. Then for each ν, N_ν is even, and for each geometric point t in T, the group G_{geom} for $\mathcal{G}_{\nu,t}$ is $Sp(N_\nu)$, and the group G_{geom} for $\mathcal{G}_{\nu,t} \oplus (\rho_{split})$ is the product group $Sp(N_\nu) \times S_{d_\nu}$.

3) (the O case) Suppose that χ has order 2, and that for every geometric point t of T, $\mathcal{F}_t$ is symplectically self–dual. Suppose also that for each ν and each geometric point t in T, there is a point β_t in $S_t - S_t \cap D_{\nu,t}$ at which $\mathcal{F}_t$ is tame, and for which

$$\mathcal{F}_t(\beta_t)/\mathcal{F}(\beta_t)^{I(\beta_t)} \text{ has odd dimension.}$$

Then for each ν, and for each geometric point t in T, the group G_{geom} for $\mathcal{G}_{\nu,t}$ is $O(N_\nu)$, and the group G_{geom} for $\mathcal{G}_{\nu,t} \oplus (\rho_{split})$ is the product group $O(N_\nu)\times S_{d_\nu}$.

4) (the strongly SO case) Suppose that χ has order 2, that the weight w is odd, that $\mathcal{F}$ is symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-w)$, and that $\mathcal{F}$ is everywhere tame. Suppose also that for each ν and each geometric point t in T, each point of S_t occurs in $D_{\nu,t}$ with even (possibly zero) multiplicity. Suppose further that for each point β_t in S_t,

$$\mathcal{F}_t(\beta_t)/\mathcal{F}(\beta_t)^{I(\beta_t)} \text{ has even dimension.}$$

Suppose further that for each finite field k, and each k–valued point t_0 of T, we have

$$\det(-\mathrm{Frob}_{k,t_0} \mid H^1(C_{t_0}\otimes\overline{k}, j_{t_0 *}\mathcal{F}_{t_0}((w+1)/2))) = 1.$$

Then for each ν, $\mathcal{G}_{\nu,t_0}((w+1)/2)$ has $G_{geom} = G_{arith} = SO(N_\nu)$, and $\mathcal{G}_{\nu,t_0}((w+1)/2) \oplus (\rho_{split})$ has $G_{geom} = G_{arith} = SO(N_\nu)\times S_{d_\nu}$.

5) (the SO/O case) Suppose that χ has order 2, that the weight w is odd, that $\mathcal{F}$ is symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-w)$, and that $\mathcal{F}$ is everywhere tame. Suppose also that for each ν and each geometric point t in T, each point of S_t occurs in $D_{\nu,t}$ with even (possibly zero) multiplicity. Suppose further that for each point β_t in S_t,

$$\mathcal{F}_t(\beta_t)/\mathcal{F}(\beta_t)^{I(\beta_t)} \text{ has even dimension.}$$

Denote by A the group homomorphism

$$A : \pi_1(T) \to \{\pm 1\}$$

given by $\det(R^1\pi_*(j_*\mathcal{F}((w+1)/2)))$, $\pi : C \to T$ the structural morphism: concretely, for each finite field k, and each k–valued point t_0 of T, we have

$$\det(-\mathrm{Frob}_{k,t_0} \mid H^1(C_{t_0}\otimes\overline{k}, j_{t_0 *}\mathcal{F}_{t_0}((w+1)/2))) = A(\mathrm{Frob}_{k,t_0}).$$

Suppose that the homomorphism A is **nontrivial.** [The case of trivial A is precisely the strongly SO case above.]

Then for each ν, $\mathcal{G}_{\nu,t_0}((w+1)/2)$ has $G_{geom} = SO(N_\nu)$, and $\mathcal{G}_{\nu,t_0}((w+1)/2) \oplus (\rho_{split})$ has $G_{geom} = SO(N_\nu)\times S_{d_\nu}$. Moreover, G_{arith} for $\mathcal{G}_{\nu,t_0}((w+1)/2) \oplus (\rho_{split})$ is equal to

$$SO(N_\nu)\times S_{d_\nu}, \text{ if } A(\mathrm{Frob}_{k,t_0}) = +1,$$
$$O(N_\nu)\times S_{d_\nu}, \text{ if } A(\mathrm{Frob}_{k,t_0}) = -1.$$

proof Statements 2), 3), 4), and 5) are fibrewise assertions, which have been proven in 5.5.1, 5.5.2, 9.5.11, 9.7.3, and 9.8.2. Statement 1) is a bit more delicate. Let us fix ν. In 5.5.1 and 5.7.1, we have proven that for each geometric point t in T, the group G_{geom} for $\mathcal{G}_{\nu,t}$ contains $SL(N_\nu)$. So either G_{geom} is the full group $GL(N_\nu)$, or it is $GL_{m_{\nu,t}}(N_\nu)$ for some integer $m_{\nu,t} \geq 1$. By Pink's semicontinuity theorem [Ka–ESDE, 8.18.2] applied to $\det(\mathcal{G}_\nu)$, there is a dense open set U_ν in T over which all the $\det(\mathcal{G}_{\nu,t})$ have the same G_{geom}, and for every t in T, G_{geom} for $\mathcal{G}_{\nu,t}$ is a subgroup of the generic G_{geom}. Given this semicontinuity, it suffices to show that for every finite field k, every k–valued point t_0 of T, $\det(\mathcal{G}_{\nu,t})$ has finite order dividing $\mu(E_\lambda)$. The point is that X_{t_0} is a smooth, geometrically connected k–scheme, and $\det(\mathcal{G}\nu,t_0)$ is an $(E_\lambda)^\times$–valued character of its entire fundamental group. But one knows [De–WeII, 1.3.4] that the restriction of any such character to the geometric fundamental group is of finite order. Since this character has values in E_λ, its finite order must be a divisor of $\mu(E_\lambda)$. QED

10.1 Definition of some measures

(10.1.1) We denote by $U_m(N)$, $USp(N)$ (if N is even), and $O(N, \mathbb{R})$ the standard compact forms of the complex groups $GL_m(N, \mathbb{C})$, $Sp(N, \mathbb{C})$, and $O(N, \mathbb{C})$ respectively, and by $U_m(N)^\#$, $USp(2N)^\#$, and $O(N, \mathbb{R})^\#$ their spaces of conjugacy classes. An agreeable feature of the $\overline{\mathbb{Q}}_\ell$–algebraic groups $GL_m(N)$, $Sp(N)$, and $O(N)$ is that for G any of these, the normalizer of G in the ambient GL(N) is $\mathbb{G}_m G$. An agreeable feature shared by the compact groups $U_m(N)$, $USp(N)$, and $O(N, \mathbb{R})$ is that in each, two elements are conjugate if and only if they have the same (reversed) characteristic polynomial $\det(1 - TA)$ in the given N–dimensional representation.

(10.1.2) Now let us put ourselves under the hypotheses and notations of Theorem 10.0.9 above.

(10.1.3) **The SL case** Fix ν. For each finite field k, and each k–valued point t of T, pick $\alpha_{\nu,k,t}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ such that $\mathcal{G}_{\nu,t}\otimes(\alpha_{\nu,k,t})^{deg}$ on $X_{\nu,t}/k$ is ι–pure of weight zero, and all its Frobenii land in $G_{geom} = GL_{m_{\nu,t}}(N_\nu)$. Then for each k–valued point x in X_t,

$$\det(1 - T\alpha_{\nu,k,t}\mathrm{Frob}_{k,t,x} \mid \mathcal{G}_\nu)$$

is the (reversed) characteristic polynomial of a unique conjugacy class

$$\theta(k, t, x, \alpha_{\nu,k,t}) \text{ in } U_{m_{\nu,t}}(N_\nu)^\#,$$

called its Frobenius conjugacy class. We define a Borel probability measure

$$\mu(k, t, \alpha_{\nu,k,t})$$

on $U_{m_{\nu,t}}(N_\nu)^\#$ to be the average, over $X_{\nu,t}(k)$, of the delta measures attached to each of these Frobenius conjugacy classes:

$$\mu(k, t, \alpha_{\nu,k,t}) := (1/\#X_{\nu,t}(k))\Sigma_{x \text{ in } X_{\nu,t}(k)}\, \delta(\theta(k, t, x, \alpha_{\nu,k,t})).$$

(10.1.4) **The Sp and O cases** Fix ν. For each finite field k, and each k–valued point t of T, pick

$\alpha_{\nu,k,t}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ such that $\mathcal{G}_{\nu,t}\otimes(\alpha_{\nu,k,t})^{\deg}$ on $X_{\nu,t}/k$ is ι–pure of weight zero, and all its Frobenii land in $G_{geom} = Sp(N_\nu)$ (resp. in $G_{geom} = O(N_\nu)$. Then for each k–valued point x in X_t,

$$\det(1 - T\alpha_{\nu,k,t}\mathrm{Frob}_{k,t,x} \mid \mathcal{G}_\nu)$$

is the (reversed) characteristic polynomial of a unique conjugacy class

$$\theta(k, t, x, \alpha_{\nu,k,t}) \text{ in } USp(N_\nu)^\# \text{ (resp. in } O(N_\nu, \mathbb{R})^\#),$$

called its Frobenius conjugacy class. We define the Borel probability measure

$$\mu(k, t, \alpha_{\nu,k,t})$$

on $Sp(N_\nu)^\#$ (resp. on $O(N_\nu, \mathbb{R})^\#$) to be the average, over $X_{\nu,t}(k)$, of the delta measures attached to each of these Frobenius conjugacy classes:

$$\mu(k, t, \alpha_{\nu,k,t}) := (1/\#X_{\nu,t}(k))\Sigma_{x \text{ in } X_{\nu,t}(k)}\ \delta(\theta(k, t, x, \alpha_{\nu,k,t})).$$

Now fix in addition a conjugacy class σ_ν in the symmetric group S_{d_ν}. The space $X_{\nu,t,\sigma_\nu\text{–split}}(k)$ is nonempty for #k sufficiently large, by 9.4.4. Whenever $X_{\nu,t,\sigma_\nu\text{–split}}(k)$ is nonempty, we define a Borel probability measure

$$\mu(k, t, \alpha_{\nu,k,t}, \sigma_\nu\text{–split})$$

on $Sp(N_\nu)^\#$ (resp. on $O(N_\nu, \mathbb{R})^\#$) to be the average, now over $X_{\nu,t,\sigma_\nu\text{–split}}(k)$, of the delta measures attached to each of these Frobenius conjugacy classes:

$$\mu(k, t, \alpha_{\nu,k,t}, \sigma_\nu\text{–split})$$
$$:= (1/\#X_{\nu,t}(k))\Sigma_{x \text{ in } X_{\nu,t,\sigma_\nu\text{–split}}(k)}\ \delta(\theta(k, t, x, \alpha_{\nu,k,t})).$$

If we are in the O case, we can further split things up according to the sign in the functional equation. Thus for each choice of sign ε, we can form the measures

$$\mu(k, t, \alpha_{\nu,k,t}, \text{sign } \varepsilon) \text{ on } O_{\text{sign } \varepsilon}(N_\nu, \mathbb{R})^\#$$

and

$$\mu(k, t, \alpha_{\nu,k,t}, \sigma_\nu\text{–split, sign } \varepsilon) \text{ on } O_{\text{sign } \varepsilon}(N_\nu, \mathbb{R})^\#,$$

respectively, as soon as $X_{\nu,t,\text{sign } \varepsilon}(k)$ and $X_{\nu,t,\sigma_\nu\text{–split,sign } \varepsilon}(k)$ are nonempty, respectively.

(10.1.5) **The strongly SO case**

Fix ν. For each finite field k, and each k–valued point t of T, pick $\alpha_{\nu,k,t}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ either choice of $\pm(\#k_i)^{(-w-1)/2}$, allowing us to define $\mathcal{G}_{\nu,t_i}((w+1)/2)$, on X_{ν,t_i}. Then $\mathcal{G}_{\nu,t}\otimes(\alpha_{\nu,k,t})^{\deg}$ on $X_{\nu,t}/k$ is ι–pure of weight zero, and all its Frobenii land in $G_{geom} = SO(N_\nu)$. For each k–valued point x in X_t, we denote by

$$\theta(k, t, x, \alpha_{\nu,k,t}) \text{ in } SO(N_\nu, \mathbb{R})^\#$$

its Frobenius conjugacy class. [In this SO case, we still have the identity

$$\det(1 - T\theta(k, t, x, \alpha_{\nu,k,t})) = \det(1 - T\alpha_{\nu,k,t}\mathrm{Frob}_{k,t,x} \mid \mathcal{G}_\nu),$$

but this identity alone only defines $\theta(\mathrm{k}, \mathrm{t}, \mathrm{x}, \alpha_{\nu,\mathrm{k},\mathrm{t}})$ as an element of $\mathrm{SO}(\mathrm{N}_\nu, \mathbb{R})$ taken up to $\mathrm{O}(\mathrm{N}_\nu, \mathbb{R})$–conjugation] We define the Borel probability measure

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}})$$

on $\mathrm{SO}(\mathrm{N}_\nu)^\#$ to be the average, over $\mathrm{X}_{\nu,\mathrm{t}}(\mathrm{k})$, of the delta measures attached to each of these Frobenius conjugacy classes:

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}}) := (1/\#\mathrm{X}_{\nu,\mathrm{t}}(\mathrm{k}))\Sigma_{\mathrm{x} \text{ in } \mathrm{X}_{\nu,\mathrm{t}}(\mathrm{k})}\, \delta(\theta(\mathrm{k}, \mathrm{t}, \mathrm{x}, \alpha_{\nu,\mathrm{k},\mathrm{t}})).$$

Now fix in addition a conjugacy class σ_ν in the symmetric group $\mathrm{S}_{\mathrm{d}_\nu}$. The space $\mathrm{X}_{\nu,\mathrm{t},\sigma_\nu\text{–split}}(\mathrm{k})$ is nonempty for #k sufficiently large by 9.4.4. Whenever $\mathrm{X}_{\nu,\mathrm{t},\sigma_\nu\text{–split}}(\mathrm{k})$ is nonempty, we define a Borel probability measure

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}}, \sigma_\nu\text{–split})$$

on $\mathrm{SO}(\mathrm{N}_\nu)^\#$to be the average, now over $\mathrm{X}_{\nu,\mathrm{t},\sigma_\nu\text{–split}}(\mathrm{k})$, of the delta measures attached to each of these Frobenius conjugacy classes:

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}}, \sigma_\nu\text{–split}) := (1/\#\mathrm{X}_{\nu,\mathrm{t}}(\mathrm{k}))\Sigma_{\mathrm{x} \text{ in } \mathrm{X}_{\nu,\mathrm{t},\sigma_\nu\text{–split}}(\mathrm{k})}\, \delta(\theta(\mathrm{k}, \mathrm{t}, \mathrm{x}, \alpha_{\nu,\mathrm{k},\mathrm{t}})).$$

(10.1.6) **The SO/O case**

Fix ν. Fix a sign $\varepsilon = \pm 1$. For each finite field k, and each k–valued point t of T with $\mathrm{A}(\mathrm{Frob}_{\mathrm{k},\mathrm{t}}) = \varepsilon$, pick $\alpha_{\nu,\mathrm{k},\mathrm{t}}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ either choice of $\pm(\#\mathrm{k}_\mathrm{i})^{(-\mathrm{w}-1)/2}$ allowing us to define $\mathcal{G}_{\nu,\mathrm{t}_\mathrm{i}}((\mathrm{w}+1)/2)$, on $\mathrm{X}_{\nu,\mathrm{t}_\mathrm{i}}$. Then $\mathcal{G}_{\nu,\mathrm{t}}\otimes(\alpha_{\nu,\mathrm{k},\mathrm{t}})^{\deg}$ on $\mathrm{X}_{\nu,\mathrm{t}}/\mathrm{k}$ is ι–pure of weight zero and orthogonally self–dual of even rank N_ν, with $\mathrm{G}_{\mathrm{geom}} = \mathrm{SO}(\mathrm{N}_\nu)$. For each k–valued point x in $\mathrm{X}_{\nu,\mathrm{t}}$, we denote by

$$\theta(\mathrm{k}, \mathrm{t}, \mathrm{x}, \alpha_{\nu,\mathrm{k},\mathrm{t}}) \text{ in } \mathrm{O}_{\mathrm{sign}\ \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$$

its Frobenius conjugacy class. We define the Borel probability measure

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}})$$

on $\mathrm{O}_{\mathrm{sign}\ \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$ to be the average, over $\mathrm{X}_{\nu,\mathrm{t}}(\mathrm{k})$, of the delta measures attached to each of these Frobenius conjugacy classes:

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}}) := (1/\#\mathrm{X}_{\nu,\mathrm{t}}(\mathrm{k}))\Sigma_{\mathrm{x} \text{ in } \mathrm{X}_{\nu,\mathrm{t}}(\mathrm{k})}\, \delta(\theta(\mathrm{k}, \mathrm{t}, \mathrm{x}, \alpha_{\nu,\mathrm{k},\mathrm{t}})).$$

Now fix in addition a conjugacy class σ_ν in the symmetric group $\mathrm{S}_{\mathrm{d}_\nu}$. The space $\mathrm{X}_{\nu,\mathrm{t},\sigma_\nu\text{–split}}(\mathrm{k})$ is nonempty for #k sufficiently large by 9.4.4. Whenever $\mathrm{X}_{\nu,\mathrm{t},\sigma_\nu\text{–split}}(\mathrm{k})$ is nonempty, we define a Borel probability measure

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}}, \sigma_\nu\text{–split})$$

on $O_{\text{sign } \varepsilon}(N_\nu, \mathbb{R})^\#$ to be the average, now over $X_{\nu,t,\sigma_\nu-\text{split}}(k)$, of the delta measures attached to each of these Frobenius conjugacy classes:

$$\mu(k, t, \alpha_{\nu,k,t}, \sigma_\nu\text{-split})$$
$$:= (1/\#X_{\nu,t}(k))\Sigma_{x \text{ in } X_{\nu,t,\sigma_\nu-\text{split}}(k)}\ \delta(\theta(k, t, x, \alpha_{\nu,k,t})).$$

Theorem 10.1.7 Hypotheses and notations as Theorem 10.0.9, we have the following results.
1) Suppose we are in the SL case. Fix ν. Suppose in addition that for **every** geometric point t in T, the lisse sheaf $\mathcal{G}_{\nu,t}$ on X_t has $G_{\text{geom}} = GL_{m_\nu}(N_\nu)$. Take any sequence of data

$$(k_i, t_i, \nu, k_i, t_i)$$

with

k_i a finite field of cardinality $\geq 4A(X_\nu/T)^2$,

t_i a k_i–valued point T,

α_{ν,k_i,t_i} in $(\overline{\mathbb{Q}}_\ell)^\times$ such that all Frobenii of $\mathcal{G}_\nu \otimes (\alpha_{k_i,t_i,\nu})^{\deg}$ land in $GL_{m_\nu}(N_\nu)$

in which $i \mapsto \#k_i$ is strictly increasing. Then the sequence of measures $\mu(k_i, t_i, \alpha_{\nu,k_i,t_i})$ on $U_{m_\nu}(N_\nu)^\#$ tends weak $*$ to (the direct image from $U_{m_\nu}(N_\nu)$ of) normalized Haar measure. In other words, for any continuous $\mathbb{C}$–valued central function f(g) on $U_{m_\nu}(N_\nu)$, we have the integration formula

$$\int_{U_{m_\nu}(N_\nu)} f(g)dg = \lim_{i\to\infty} \int_{U_{m_\nu}(N_\nu)} f(g)d\mu(k_i, t_i, \alpha_{\nu,k_i,t_i})$$
$$= \lim_{i\to\infty} (1/\#X_{\nu,t_i}(k_i))\Sigma_{x \text{ in } X_{\nu,t_i}(k_i)}\ f(\theta(k_i, t_i, x, \alpha_{\nu,k_i,t_i})).$$

2) Suppose we are in the Sp or O case. Fix ν, and fix a conjugacy class in the symmetric group S_{d_ν}. Take any sequence of data

$$(k_i, t_i, \alpha_{\nu,k_i,t_i})$$

with

k_i a finite field, $\#k_i > \text{Max}(4A(X_\nu/T)^2, 4C(X_\nu/T, \mathcal{S}_\Gamma)^2(\#\Gamma)^4)$

t_i a k_i–valued point T,

α_{ν,k_i,t_i} in $(\overline{\mathbb{Q}}_\ell)^\times$ such that all Frobenii of $\mathcal{G}_\nu \otimes (\alpha_{k_i,t_i,\nu})^{\deg}$ land in $Sp(N_\nu)$ (resp. in $O(N_\nu)$)

in which $i \mapsto \#k_i$ is strictly increasing. Then the two sequences of measures $\mu(k_i, t_i, \alpha_{\nu,k_i,t_i})$ and $\mu(k_i, t_i, \alpha_{\nu,k_i,t_i}, \sigma_\nu\text{-split})$ on $USp(N_\nu)^\#$ (resp. on $O(N_\nu, \mathbb{R})^\#$) each tend weak $*$ to (the direct image from $USp(N_\nu)$ (resp. from $O(N_\nu, \mathbb{R})$) of) normalized Haar measure. In the O case, the sequences of measures

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}}, \text{sign } \varepsilon) \text{ on } \mathrm{O}_{\text{sign } \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$$

and

$$\mu(\mathrm{k}, \mathrm{t}, \alpha_{\nu,\mathrm{k},\mathrm{t}}, \sigma_\nu\text{–split}, \text{sign } \varepsilon) \text{ on } \mathrm{O}_{\text{sign } \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$$

each tend weak $*$ to Haar measure on $\mathrm{O}_{\text{sign } \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$ normalized now to give $\mathrm{O}_{\text{sign } \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$ total mass one.

3) Suppose we are in the strongly SO case. Fix ν, and fix a conjugacy class in the symmetric group $\mathrm{S}_{\mathrm{d}_\nu}$. Take any sequence of data

$$(\mathrm{k}_\mathrm{i}, \mathrm{t}_\mathrm{i}, {}_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}})$$

with

k_i a finite field, $\#\mathrm{k}_\mathrm{i} > \mathrm{Max}(4\mathrm{A}(\mathrm{X}_\nu/\mathrm{T})^2, 4\mathrm{C}(\mathrm{X}_\nu/\mathrm{T}, \mathcal{S}_\Gamma)^2(\#\Gamma)^4)$

t_i a k_i–valued point T,

$\alpha_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ either choice of $\pm(\#\mathrm{k}_\mathrm{i})^{(-\mathrm{w}-1)/2}$,

in which $\mathrm{i} \mapsto \#\mathrm{k}_\mathrm{i}$ is strictly increasing. Then the two sequences of measures $\mu(\mathrm{k}_\mathrm{i}, \mathrm{ti}, \alpha_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}})$ and $\mu(\mathrm{k}_\mathrm{i}, \mathrm{ti}, \alpha_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}}, \sigma_\nu\text{–split})$ on $\mathrm{SO}(\mathrm{N}_\nu)^\#$ each tend weak $*$ to (the direct image from $\mathrm{SO}(\mathrm{N}_\nu)$ of) normalized Haar measure.

4) Suppose we are in the SO/O case. Fix ν, fix a sign $\varepsilon = \pm 1$, and fix a conjugacy class in the symmetric group $\mathrm{S}_{\mathrm{d}_\nu}$. Take any sequence of data

$$(\mathrm{k}_\mathrm{i}, \mathrm{t}_\mathrm{i}, {}_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}})$$

with

k_i a finite field, $\#\mathrm{k}_\mathrm{i} > \mathrm{Max}(4\mathrm{A}(\mathrm{X}_\nu/\mathrm{T})^2, 4\mathrm{C}(\mathrm{X}_\nu/\mathrm{T}, \mathcal{S}_\Gamma)^2(\#\Gamma)^4)$

t_i a k_i–valued point T such that $\mathrm{A}(\mathrm{Frob}_{\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}}) = \varepsilon$,

$\alpha_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}}$ in $(\overline{\mathbb{Q}}_\ell)^\times$ either choice of $\pm(\#\mathrm{k}_\mathrm{i})^{(-\mathrm{w}-1)/2}$,

in which $\mathrm{i} \mapsto \#\mathrm{k}_\mathrm{i}$ is strictly increasing. Then the two sequences of measures $\mu(\mathrm{k}_\mathrm{i}, \mathrm{ti}, \alpha_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}})$ and $\mu(\mathrm{k}_\mathrm{i}, \mathrm{ti}, \alpha_{\nu,\mathrm{k}_\mathrm{i},\mathrm{t}_\mathrm{i}}, \sigma_\nu\text{–split})$ on $\mathrm{O}_{\text{sign } \varepsilon}(\mathrm{N}_\nu)^\#$ each tend weak $*$ to Haar measure on $\mathrm{O}_{\text{sign } \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$ normalized to give $\mathrm{O}_{\text{sign } \varepsilon}(\mathrm{N}_\nu, \mathbb{R})^\#$ total mass one.

proof Assertion 1) is a restatement of 9.5.11, with Γ there taken to be the trivial group {e}. Assertion 2) is a restatement of Theorem 9.10.3 and Corollary 9.10.7. Assertions 3) and 4) are restatements of Theorems 9.10.4 and 9.10.8, respectively. QED

10.2 Some basic examples of data (C/T, S, $\mathcal{F}$, D_ν's) where all the hypotheses above are satisfied

10.2.1 **SL examples**

(10.2.1.1) This first example is the "universal" form of the situation considered in Theorem 7.9.1. Fix an integer $n \geq 3$, a prime ℓ, a $(\overline{\mathbb{Q}}_\ell)^\times$–valued character χ of order n of the group $\mu_n(\mathbb{Z}[1/n, \zeta_n])$, and an integer $g \geq 2$. Denote by $\mathcal{M}_{g,3K}$ the moduli space of genus g curves with a level 3K structure, cf. [Ka–Sa, RMFEM, 10.6], and denote by $\mathcal{C}_{univ}/\mathcal{M}_{g,3K}$ the universal genus g curve with level 3K structure. For each integer $m \geq 1$, denote by $\mathcal{M}_{g,3K,m} := (\mathcal{C}_{univ}/\mathcal{M}_{g,3K})^m$ the m–fold fibre product of $\mathcal{C}_{univ}$ with itself over $\mathcal{M}_{g,3K}$. This space $\mathcal{M}_{g,3K,m}$ is the moduli space of genus g curves with both a level 3K structure and with an ordered list of m points $P_1, ..., P_m$, not necessarily distinct. Denote by $\mathcal{M}_{g,3K,m\ dist}$ the open set in $\mathcal{M}_{g,3K,m}$ where, for all $i \neq j$, P_i and P_j are disjoint. Thus $\mathcal{M}_{g,3K,m\ dist}$ is the moduli spaces of curves of genus g with both a 3K structure and with an ordered list of m distinct points $P_1, ..., P_m$. Denote by $\mathcal{C}_{univ,m}/\mathcal{M}_{g,3K,m\ dist}$ the universal curve. We take

$$T := \mathcal{M}_{g,3K,m\ dist} \times_{\mathbb{Z}} \mathbb{Z}[1/\ell n, \zeta_n],$$

and we take C/T to be universal curve $\mathcal{C}_{univ,m} \times_{\mathbb{Z}} \mathbb{Z}[1/\ell n, \zeta_n]$. We take S to be empty, and $\mathcal{F}$ to be the constant sheaf $\overline{\mathbb{Q}}_\ell$. We take D_ν to be any divisor of the form $\Sigma_{i=1 \text{ to } n} a_{i,\nu} P_i$ where the P_i are the tautological points, and where the $a_{i,\nu}$ are nonnegative integers with $\Sigma_i a_{i,\nu} \geq 4g+4$ and increasing with ν. If n is 2×(odd), require further that each $a_{i,\nu}$ is either odd or divisible by n. In this case, the common value of G_{geom} for $\mathrm{Twist}_{\chi,C/T,D_\nu}(\overline{\mathbb{Q}}_\ell)$ on all geometric fibres of X_ν/T is $GL_\mu(N_\nu)$, where μ is the order of the character $\chi \times \chi_2$. [So μ is 2n if n is odd, μ is n/2 if n is 2×(odd), and μ is n if $n \equiv 0 \bmod 4$).]

(10.2.2) **Sp and O examples** In all these examples, we take n=2. We begin with three elliptic curve examples.

(10.2.2.1) Take n=2, $T = \mathrm{Spec}(\mathbb{Z}[1/2\ell])$, $C/T = \mathbb{P}^1/T$, S is $\{0,1,\infty\}$. The open curve C – S is thus $\mathrm{Spec}(\mathbb{Z}[1/2\ell, \lambda, 1/\lambda(\lambda-1)]$. Take $\mathcal{F}_1$ to be $R^1\pi_!\overline{\mathbb{Q}}_\ell$ for π the structural morphism of the Legendre family Leg/(C–S) of elliptic curves

$$y^2 = x(x-1)(x-\lambda).$$

Then $\mathcal{F}_1$ is lisse of rank 2 on C–S, pure of weight one, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-1)$. Along the sections 0 and 1 of C/T, $\mathcal{F}$ has unipotent nontrivial local monodromy. Along the section ∞, its monodromy is (the quadratic character)⊗(unipotent nontrivial). For each integer $n \geq 1$, take $\mathcal{F}_n := \mathrm{Sym}^n(\mathcal{F}_1)$. Thus $\mathcal{F}_n$ is lisse of rank n+1, pure of weight n, and autodual toward $\overline{\mathbb{Q}}_\ell(-n)$, by an autoduality which is symplectic for odd n, and orthogonal for even n. The local monodromy along the sections 0 and 1 is a single unipotent Jordan block. The local monodromy along ∞ is a single unipotent Jordan block for n even, and

(the quadratic character)$\otimes$(a single unipotent Jordan block)

for n odd. We take for D_ν the divisor $d_\nu\infty$. So here we are performing quadratic twists of the $\mathcal{F}_n$'s by polynomials in λ of degree d_ν which have d_ν distinct zeroes, none of which is 0 or 1. For n odd (resp. for n even), the sheaves $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_n)$ have G_{geom} the full orthogonal group (resp. the full symplectic group) on each geometric fibre of X_ν/T.

(10.2.2.2) In a similar vein, we might take some level $m \geq 3$, and then take

$$T = \mathrm{Spec}(\mathbb{Z}[1/m\ell, \zeta_m]),$$

and C/T the compactified moduli space $\overline{\mathcal{M}}_m[1/m\ell]$ of elliptic curves with level m structure of determinant ζ_m over $\mathbb{Z}[1/m\ell, \zeta_m]$–schemes, S the cusps. We take

$$\pi : \mathcal{E}_{univ,m} \to \mathcal{M}_m[1/m\ell] = C-S$$

the universal curve, and $\mathcal{F}_1 := R^1\pi_!\overline{\mathbb{Q}}_\ell$. Once again $\mathcal{F}_1$ is lisse of rank 2 on C−S, pure of weight one, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-1)$. Its local monodromy along each cusp is unipotent nontrivial. For each integer $n \geq 1$, take $\mathcal{F}_n := \mathrm{Sym}^n(\mathcal{F}_1)$. Thus $\mathcal{F}_n$ is lisse of rank n+1, pure of weight n, and autodual toward $\overline{\mathbb{Q}}_\ell(-n)$, by an autoduality which is symplectic for odd n, and orthogonal for even n. The local monodromy along each cusp is a single unipotent Jordan block. Take the D_ν's to be divisors concentrated at the cusps. When n is odd, $\mathcal{F}_n$ is symplectic. In this case, we must require that each divisor D_ν omits at least one cusp (so that there is a finite singularity where the drop is of odd dimension, which in turn will insure that for each t, $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_n)$ has Ggeom the full orthogonal group. When n is even, $\mathcal{F}_n$ is orthogonal, and for each t in T the sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_n)$ has Ggeom the full symplectic group.

(10.2.2.3) Take K to be an absolutely finitely generated subfield of $\mathbb{C}$, C_K/K a proper smooth geometrically connected curve over K, with function field L/K, and E/L an elliptic curve over L. We make one hypothesis on E/L, namely that at K–valued point P_K of C_K, i.e., at some discrete valuation of L/K with residue field K, E/L has multiplicative reduction. We can find a dense open set U_K in C_K and an elliptic curve E_K/U_K whose generic fibre is E/L. [Concretely, take the Néron model $\mathcal{E}_K/C_K$ of E/L and take U_K to be the open set of C_K over which the Néron model is an elliptic curve.] Fix a prime number ℓ. We can then find

a) a subring R of K in which 2ℓ in invertible, which is finitely generated as a $\mathbb{Z}[1/2\ell]$–algebra and which is smooth over $\mathbb{Z}$,

b) a proper smooth curve C/R with geometrically connected fibres, and an R–valued point P in C(R) which extends P_K,

c) an effective divisor S in C which is finite etale over R, contains P, and whose open complement $U := C - S$ has generic fibre U_K/K,

d) and an elliptic curve $\pi : E \to U$ which extends E_K/U_K.

We take T := Spec(R), C/T and S/T as above, and for lisse sheaf $\mathcal{F}_1$ on U we take $R^1\pi_*\overline{\mathbb{Q}}_\ell$. We take for the D_ν effective divisors whose supports $(D_\nu)^{red}$ lie in S–P (this insures that on each geometric fibre of $(C - D_\nu)/T$, there is a point (namely P) at which $\mathcal{F}_1$ has nontrivial unipotent monodromy. We take $\mathcal{F}_n := \mathrm{Sym}^n(\mathcal{F}_1)$, and proceed as in examples 1) and 2) above.

(10.2.2.4) We now give two examples involving hyperelliptic curves.

(10.2.2.5) Fix an integer $m \geq 2$, and take T to be the open set in $\mathbb{A}^m \times \mathbb{G}_m/\mathbb{Z}[1/2\ell]$, with coordinates $a_0, a_1, \ldots, a_m$ over which the degree m polynomial in one variable

$$f(x) := \Sigma_i a_i x^i$$

has invertible discriminant Δ (i.e., has d distinct roots). Take C/T to be $\mathbb{P}^1/T$, S to be

{zeroes of f}, if m is even

$\{\infty\} \cup$ {zeroes of f}, if m is odd.

Take $\mathcal{F}_0$ on $\mathbb{P}^1 - S$ to be $\mathcal{L}_{\chi_2(f(x))}$, which is orthogonally self–dual, and pure of weight zero. Take D_ν to be the divisor $d_\nu\infty$. Then for each t in T, $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_0)$ has G_{geom} the full symplectic group. Concretely, for fixed t in T, corresponding to a numerical choice of polynomial f, $X_{\nu,t}$ is the space of polynomials p(x) of degree d_ν with all distinct roots and with

$$\text{g.c.d.}(p(x), f(x)) = 1.$$

Over this space we are looking at the family of hyperelliptic curves

$$y^2 = f(x)p(x),$$

parameterized by the polynomial p(x), and our $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_0)$ is the H^1 along the fibres in this family.

(10.2.2.6) Notations as in 10.2.2.5 above, take $\mathcal{F}_{0,!}$ to be the extension by zero to $\mathbb{A}^1$ of (the restriction to $\mathbb{A}^1 - \mathbb{A}^1 \cap S$ of) $\mathcal{F}_0$. Define $\mathcal{F}_1$ on $\mathbb{A}^1 - \mathbb{A}^1 \cap S$ to be the lisse sheaf which is the restriction from $\mathbb{A}^1$ of the middle convolution of $\mathcal{F}_{0,!}$ with $\mathcal{L}_{\chi_2}$ on $\mathbb{A}^1$. The rank of $\mathcal{F}_1$ is m if m is even, m–1 if m is odd. For each t $\mathcal{F}_{1,t}$ has G_{geom} the full symplectic group Sp(m) if m is even, Sp(m–1) if m is odd. Local monodromy of $\mathcal{F}_1$ along each of the m zeroes of f is a unipotent pseudoreflection (transvection). Local monodromy along ∞ is

(the quadratic character)$\otimes$(a unipotent pseudoreflection)

if m is even. If m is odd, local monodromy along ∞ is scalar, the quadratic character. Take D_ν to be the divisor $d_\nu\infty$. Then for each t in T, $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_1)$ has G_{geom} the full orthogonal group.

Here is a more geometric description of the sheaf $\mathcal{F}_1$. Over T as in 4) above, consider

$(\mathbb{A}^1 - \mathbb{A}^1 \cap S)/T$ with parameter λ, i.e., consider $\mathbb{A}^1[1/f(\lambda)]/T$. Over this $\mathbb{A}^1[1/f(\lambda)]/T$, we have the complete nonsingular model $\pi: \mathcal{C} \to \mathbb{A}^1[1/f(\lambda)]/T$ of the hyperelliptic curve with equation

$$y^2 = f(x)(\lambda - x).$$

Then $\mathcal{F}_1$ is the sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell$ on $\mathbb{A}^1[1/f(\lambda)]/T$. The interpretation of the twist sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_1)$ is this. For fixed t in T, corresponding to a numerical choice of polynomial f, $X_{\nu,t}$ is the space of polynomials $p(\lambda)$ of degree d_ν with all distinct roots and with

$$\text{g.c.d.}(p(\lambda), f(\lambda)) = 1.$$

The twist sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_1)$ on $X_{\nu,t}$ gives the L–functions of the quadratic twists, by polynomials of degree d_ν in λ with all distinct roots and with g.c.d.$(p(\lambda), f(\lambda)) = 1$, of the Jacobian of the hyperelliptic curve $y^2 = f(x)(\lambda - x)$, viewed as curve over the λ–line.

(10.2.3) **Strongly SO examples**

(10.2.3.1) Take n=2, $T = \mathrm{Spec}(\mathbb{Z}[i, 1/2\ell])$, $C/T = \mathbb{P}^1/T$, $S = \{0,1,\infty\}$. The open curve $C - S$ is thus $\mathrm{Spec}(\mathbb{Z}[1/2\ell, \lambda, 1/\lambda(\lambda-1)])$. Take $\mathcal{F}_1$ to be $R^1\pi_!\overline{\mathbb{Q}}_\ell$ for π the structural morphism of the **twisted** Legendre family of elliptic curves

$$y^2 = \lambda(\lambda-1)x(x-1)(x-\lambda).$$

Then $\mathcal{F}_1$ is lisse of rank 2 on C–S, pure of weight one, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-1)$. Along the sections 0, 1, and ∞ of C/T, the local monodromy of $\mathcal{F}$ is

(the quadratic character)$\otimes$(unipotent nontrivial).

For each **odd** integer $m \geq 1$, take $\mathcal{F}_m := \mathrm{Sym}^m(\mathcal{F}_1)$. Thus $\mathcal{F}_m$ is lisse of even rank m+1, pure of weight m, and orthogonally self–dual toward $\overline{\mathbb{Q}}_\ell(-m)$. Its local monodromy along the sections 0, 1, ∞ is

(the quadratic character)$\otimes$(a single unipotent Jordan block).

Suppose each d_ν is **even**, and take for D_ν the divisor $d_\nu\infty$. So here we are performing quadratic twists of the $\mathcal{F}_m$'s by polynomials in λ of even degree d_ν which have d_ν distinct zeroes, none of which is 0 or 1. For each odd m, $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ has rank $(m+1)(d_\nu + 1)$. By 8.5.7, for $\nu >> 0$, G_{geom} for $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_n)$ is the group $SO((m+1)(d_\nu + 1))$ on each geometric fibre of X_ν/T. By 8.9.2, for each finite field k and each k–valued point of T, the sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ on $X_\nu\otimes_T k$ has $G_{arith} = SO((m+1)(d_\nu + 1))$. Indeed, if $T = \mathrm{Spec}(\mathbb{Z}[i, 1/2\ell])$ admits a k–valued point, then k has odd characteristic not ℓ, and k contains a primitive fourth root of unity. Thus $\#k \equiv 1 \bmod 4$, and we apply 8.9.2.

(10.2.3.2) Take n=2, $T = \mathrm{Spec}(\mathbb{Z}[1/2\ell])$, $C/T = \mathbb{P}^1/T$, $S = \{0,1,\infty\}$. For each positive integer $m \equiv 3 \bmod 4$, take $\mathcal{F}_m$ from the example 10.2.3.1 above, and take the D_ν as in that example. By 8.5.7 and

8.9.2, for each finite field k of odd characteristic not ℓ, the sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ on $X_\nu\otimes_T k$ has $G_{geom} = G_{arith}$ = the group $SO((m+1)(d_\nu + 1))$.

(10.2.3.3) Take n=2, T = Spec($\mathbb{Z}$[i, 1/6ℓ, Δ, 1/Δ]), C/T = E_Δ/T the elliptic curve whose affine equation in (g_2, g_3)–space is

$$(g_2)^3 - 27(g_3)^2 = \Delta,$$

S = {∞}, the origin on E_Δ. On C – S, take $\mathcal{F}_1$ to be $R^1\pi_!\overline{\mathbb{Q}}_\ell$ for π the structural morphism of the universal family of elliptic curves with differential (E, ω) with discriminant Δ

$$y^2 = 4x^3 - g_2x - g_3.$$

Then $\mathcal{F}_1$ is lisse of rank 2 on C–S, pure of weight one, and symplectically self–dual toward $\overline{\mathbb{Q}}_\ell(-1)$. Along the identity section ∞ of C/T, the local monodromy of $\mathcal{F}$ is

(the quadratic character)$\otimes$(unipotent nontrivial).

For each **odd** integer m $\geq$ 1, take $\mathcal{F}_m := \mathrm{Sym}^m(\mathcal{F}_1)$. Thus $\mathcal{F}_m$ is lisse of even rank m+1, pure of weight m, and orthogonally self–dual toward $\overline{\mathbb{Q}}_\ell(-m)$. Its local monodromy along the identity section ∞ is

(the quadratic character)$\otimes$(a single unipotent Jordan block).

Suppose each d_ν is **even**, and take for D_ν the divisor $d_\nu\infty$. So here we are performing quadratic twists of the $\mathcal{F}_m$'s by polynomials in x and y which have a pole at ∞ of even degree d_ν and which have d_ν distinct zeroes. For each odd m, $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ has rank $(m+1)(d_\nu + 1)$. By 8.5.7, for $\nu >> 0$, G_{geom} for $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_n)$ is the group $SO((m+1)(d_\nu + 1))$ on each geometric fibre of X_ν/T. By 8.10.6, for each finite field k and each k–valued point of T, the sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ on $X_\nu\otimes_T k$ has $G_{arith} = SO((m+1)(d_\nu + 1))$. Indeed, if T = Spec($\mathbb{Z}$[i, 1/6$\ell$]) admits a k–valued point, then k has characteristic prime to 6ℓ, and k contains a primitive fourth root of unity. Thus #k $\equiv$ 1 mod 4, and we apply 8.10.6.

(10.2.3.4) Take n=2, T = Spec($\mathbb{Z}$[1/6ℓ, Δ, 1/Δ]), C/T = E_Δ/T, S = {∞}, the origin on E_Δ. For each positive integer m $\equiv$ 3 mod 4, take $\mathcal{F}_m$ from the example 10.2.3.3 above, and take the D_ν as in that example. By 8.5.7 and 8.10.6, for each finite field k of characteristic prime to 6ℓ, the sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ on $X_\nu\otimes_T k$ has $G_{geom} = G_{arith}$ = the group $SO((m+1)(d_\nu + 1))$.

(10.2.4) **SO/O examples**

(10.2.4.1) Take n=2, T = Spec($\mathbb{Z}$[1/2ℓ]), C/T = $\mathbb{P}^1$/T, S = {0,1,∞}. For each positive integer m $\equiv$ 1 mod 4, take $\mathcal{F}_m$ from the example 10.2.3.1 above, and take the D_ν as in that example. By 8.5.7 and 8.9.2, for each finite field k of odd characteristic not ℓ, the sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ on $X_\nu\otimes_T k$

has $G_{geom} = SO((m+1)(d_\nu + 1))$. If $\#k \equiv 1 \bmod 4$, then $G_{arith} = G_{geom} = SO((m+1)(d_\nu + 1))$, but if $\#k \equiv 3 \bmod 4$, then G_{arith} is $O((m+1)(d_\nu + 1))$.

(10.2.4.2) Take n=2, $T = \mathrm{Spec}(\mathbb{Z}[1/6\ell, \Delta, 1/\Delta])$, $C/T = E_\Delta/T$, $S = \{\infty\}$, the origin on E_Δ. For each positive integer $m \equiv 1 \bmod 4$, take $\mathcal{F}_m$ from the example 10.2.3.3 above, and take the D_ν as in that example. By 8.5.7 and 8.10.6, for each finite field k of characteristic prime to 6ℓ, the sheaf $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_m)$ on $X_\nu \otimes_T k$ has $G_{geom} = SO((m+1)(d_\nu + 1))$. If $\#k \equiv 1 \bmod 4$, then $G_{arith} = G_{geom} = SO((m+1)(d_\nu + 1))$, but if $\#k \equiv 3 \bmod 4$, then G_{arith} is $O((m+1)(d_\nu + 1))$.

(10.2.5) **More SL examples**

(10.2.5.1) We take $n \geq 3$ **odd**, $\chi : \mu_n(\mathbb{Z}[1/n\ell, \zeta_n]) \to (\overline{\mathbb{Q}}_\ell)^\times$ a character of order n. Pick an integer $m \geq 2$. Take T to be the open set in $\mathbb{A}^m \times \mathbb{G}_m/\mathbb{Z}[1/2n\ell, \zeta_n]$, with coordinates $a_0, a_1, \ldots, a_m$ over which the degree m polynomial in one variable

$$f(x) := \Sigma_i a_i x^i$$

has invertible discriminant Δ (i.e., has d distinct roots). Take C/T to be $\mathbb{P}^1/T$, S to be

{zeroes of f}, if $m \equiv 0 \bmod n$,

$\{\infty\} \cup$ {zeroes of f} if m is nonzero mod n.

Take $\mathcal{F}_0$ on $\mathbb{P}^1 - S$ to be $\mathcal{L}_{\chi(f(x))}$. Take D_ν to be the divisor $d_\nu\infty$.

Concretely, for fixed t in T, corresponding to a numerical choice of polynomial f, $X_{\nu,t}$ is the space of polynomials p(x) of degree d_ν with all distinct roots and with g.c.d.(p(x), f(x)) = 1. Over this space we are looking at the family of curves

$$y^n = f(x)p(x),$$

parameterized by the polynomial p(x). The group μ_n acts (by moving y) on this family, and our $\mathrm{Twist}_{\chi_2,C/T,D_\nu}(\mathcal{F}_0)$ is the χ–component of the H^1 along the fibres in this family.

We claim that for each t in T, $\mathrm{Twist}_{\chi,C/T,D_\nu}(\mathcal{F}_0)$ has G_{geom} the group $GL_{2n}(N_\nu)$. By Pink's semicontinuity result [Ka–ESDE, 8.18.2], it suffices to check at t (lying over) a finite field valued point of T. So we may assume that T is Spec(k) with k a finite field. We must show that $\det(\mathcal{G}_\nu)$ is geometrically of order 2n. Because we took n to be odd, 2n is the number of roots of unity in the field $\mathbb{Q}(\chi)$. We use the "compatible system over $\mathbb{Q}(\chi)$" argument of [Ka–ACT, the "trivial" part of the proof of 5.2 bis], already used in 7.9.2, 7.9.3, and 7.10.2, to see that $\det(\mathcal{G}_\nu)^{\otimes 2n}$ is trivial. We use a one–parameter family of twists of the form $t \mapsto (t - p_1(x))p_2(x)$ to get a curve in X_ν along which $\mathcal{G}_\nu$ has some local monodromies which are pseudoreflections of determinant $\chi \times \chi_2$, cf. 5.4.9. So already $\det(\mathcal{G}_\nu)$ has geometric order at least 2n along this curve, and hence $\det(\mathcal{G}_\nu)$ is geometrically of order 2n on X_ν, as required.

(10.2.5.2) Notations as in 10.2.5.1 above, take $\mathcal{F}_{0,!}$ to be the extension by zero to $\mathbb{A}^1$ of (the restriction to $\mathbb{A}^1 - \mathbb{A}^1 \cap S$ of) $\mathcal{F}_0$. Define $\mathcal{F}_1$ on $\mathbb{A}^1 - \mathbb{A}^1 \cap S$ to be the lisse sheaf which is the restriction from $\mathbb{A}^1$ of the middle convolution of $\mathcal{F}_{0,!}$ with $\mathcal{L}_\chi$ on $\mathbb{A}^1$. The rank of $\mathcal{F}_1$ is m unless $m \equiv -1 \bmod n$, in which case the rank is $m-1$. Local monodromy of $\mathcal{F}_1$ along each of the m zeroes of f is a pseudoreflection of determinant χ^2. Local monodromy along ∞ is

$$\chi \otimes (\text{a pseudoreflection of determinant } \chi^m)$$

unless $m \equiv -1 \bmod n$. If $m \equiv -1 \bmod n$, local monodromy along ∞ is scalar, the character χ. For each t $\mathcal{F}_{1,t}$ has G_{geom} the group $GL_n(m)$ unless $m \equiv -1 \bmod n$, and in that case G_{geom} is $GL_n(m-1)$. To see this, use the fact that G_{geom} contains SL, and then use the local monodromy information to compute the tame sheaf $\det(\mathcal{F}_{1,t})$.

Take D_ν to be the divisor $d_\nu\infty$. For each t in T, $\mathcal{G}_\nu := \mathrm{Twist}_{\chi,C/T,D_\nu}(\mathcal{F}_1)$ has G_{geom} the group $GL_{2n}(N_\nu)$. One sees this by using the fact that G_{geom} contains SL, and then computing the geometric order of $\det(\mathcal{G}_{\nu,t})$ at finite field valued points t of T by the argument used in the previous example. The compatible system argument again shows that $\det(\mathcal{G}_{\nu,t})^{\otimes 2n}$ is geometrically trivial. The same sort of one–parameter family of twists as used above again produces a curve in $X_{\nu,t}$ along which $\mathcal{G}_{\nu,t}$ has some local monodromies which are pseudoreflections of determinant $\chi \times \chi_2$, and one concludes exactly as above.

10.3 Applications to average rank

Theorem 10.3.1 Suppose we have $\mathcal{F}$ on (C–S)/T satisfying all the hypotheses of Theorem 10.0.9, part 3). Fix ν, and fix a conjugacy class σ_ν in the symmetric group S_{d_ν}. Take any sequence of data

$$(k_i, t_i, \alpha_{\nu,k_i,t_i})$$

with

k_i a finite field, $\#k_i > \mathrm{Max}(4A(X_\nu/T)^2, 4C(X_\nu/T, \mathcal{S}_{S_{d_\nu}})^2(d_\nu!)^4)$

t_i a k_i–valued point T,

α_{ν,k_i,t_i} in $(\overline{\mathbb{Q}}_\ell)^\times$ such that all Frobenii of $\mathcal{G}_\nu \otimes (\alpha_{k_i,t_i,\nu})^{\deg}$ land in $O(N_\nu)$, i.e., α_{ν,k_i,t_i} is **any** choice of a square root of $(\#k_i)^{-w-1}$, allowing us to define $\mathcal{G}_{\nu,t_i}((w+1)/2)$, on X_{ν,t_i},

in which $i \mapsto \#k_i$ is strictly increasing. Then we have the following table of limit formulas. In these tables, the number in the third column is the limit, as $i \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1/2,
$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{quad,\, an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{geom,\, an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1.

More precisely, for each finite extension E/k, and each value of $\varepsilon = \pm 1$, denote by $X_{\nu,t_i,\sigma_\nu\text{-split, sign }\varepsilon}(k_i)$ the subset of $X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$ consisting of those points f in $X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$ such that

$$\det(-\alpha_{\nu,k_i,t_i}\mathrm{Frob}_{k_i,f} \mid \mathcal{G}_\nu) = \varepsilon.$$

Then we have the following table of limit formulas. In these tables, the number in the third column is the limit, as $i \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

If N_ν is even:

$X_{\nu,t_i,\sigma_\nu\text{-split, sign }-}(k_i)$	$\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
$X_{\nu,t_i,\sigma_\nu\text{-split, sign }+}(k_i)$	$\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0,
$X_{\nu,t_i,\sigma_\nu\text{-split, sign }-}(k_i)$	$\mathrm{rank}_{quad,\, an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	2,
$X_{\nu,t_i,\sigma_\nu\text{-split, sign }+}(k_i)$	$\mathrm{rank}_{quad,\, an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0,
$X_{\nu,t_i,\sigma_\nu\text{-split, sign }-}(k_i)$	$\mathrm{rank}_{geom,\, an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	2,
$X_{\nu,t_i,\sigma_\nu\text{-split, sign }+}(k_i)$	$\mathrm{rank}_{geom,\, an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0.

If N_ν is odd:

X_{ν,t_i,σ_ν}–split, sign $-^{(k_i)}$	$\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
X_{ν,t_i,σ_ν}–split, sign $+^{(k_i)}$	$\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0,
X_{ν,t_i,σ_ν}–split, sign $-^{(k_i)}$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
X_{ν,t_i,σ_ν}–split, sign $+^{(k_i)}$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
X_{ν,t_i,σ_ν}–split, sign $-^{(k_i)}$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
X_{ν,t_i,σ_ν}–split, sign $+^{(k_i)}$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1.

proof Immediate from Theorem 10.1.7, part 2), and the proof of 8.3.3. QED

Remark 10.3.2 Notice that $\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$ is defined as the order of vanishing at T=1 of $\det(1 - TF \mid \mathcal{G}_{\nu,t_i,f}((w+1)/2))$, and that $\mathcal{G}_{\nu,t_i}((w+1)/2)$ was **defined** to be $\mathcal{G}_{\nu,t_i}\otimes(\alpha_{\nu,k_i,t_i})^{\deg}$. In other words, the analytic rank in question is the order of vanishing of

$$\det(1 - T\,\mathrm{Frob}_{k_i,f} \mid \mathcal{G}_\nu)$$

at the point $T = \alpha_{\nu,k_i,t_i}$. So this notion **depends** on **which choice** of square root of $(\#k_i)^{-w-1}$ we take for α_{ν,k_i,t_i}. The quadratic and geometric analytic ranks do not depend on this choice. The reader may at first be disturbed that our results on average analytic rank apply equally to order of vanishing at the two different points, but there is no contradiction. On the compact group $O(N_\nu, \mathbb{R})$, $A \mapsto -A$ is a (measure–preserving) involution which interchanges the functions

$$A \mapsto \text{order of vanishing of } \det(1-TA) \text{ at } T=1$$

and

$$A \mapsto \text{order of vanishing of } \det(1-TA) \text{ at } T=-1.$$

In the case when $\mathcal{F}$ arises as the H^1 along the fibres of a family of abelian varieties, its weight w is 1, and it is the choice $(\#k_i)^{-1}$ of square root of $(\#k_i)^{-2}$ which must be taken in defining $\mathcal{G}_\nu(1)$ in the Birch and Swinnerton–Dyer conjecture. This problem did not arise in our earlier discussion 8.1.1 of average rank over a fixed finite field k, because earlier (7.0.9) we chose a square root α_k of #k, and agreed to use powers of α_k whenever we needed square roots of integer powers of #k.

Theorem 10.3.3 Suppose we have $\mathcal{F}$ on (C–S)/T satisfying all the hypotheses of Theorem 10.0.9, part 4). Then $\mathcal{G}_\nu$ is orthogonally self–dual toward $\overline{\mathbb{Q}}_\ell(-w-1)$. Fix ν, and fix a conjugacy class σ_ν

in the symmetric group S_{d_ν}. Take any sequence of data

$$(k_i, t_i, \alpha_{\nu,k_i,t_i})$$

with

k_i a finite field, $\#k_i > \mathrm{Max}(4A(X_\nu/T)^2, 4C(X_\nu/T, \mathcal{S}_{S_{d_\nu}})^2(d_\nu!)^4)$

t_i a k_i–valued point T,

α_{ν,k_i,t_i} in $(\overline{\mathbb{Q}}_\ell)^\times$ is either choice of $\pm(\#k_i)^{(-w-1)/2}$,

in which $i \mapsto \#k_i$ is strictly increasing. Each set $X_{\nu,t_i,\mathrm{sign}\,-}$ is empty. We have the following table of limit formulas. In these tables, the number in the third column is the limit, as $i \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

$X_{\nu,t_i,\sigma_\nu\text{–split}}(k_i)$	$\mathrm{rank}_{\mathrm{an}}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0,
$X_{\nu,t_i,\sigma_\nu\text{–split}}(k_i)$	$\mathrm{rank}_{\mathrm{quad,\,an}}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0,
$X_{\nu,t_i,\sigma_\nu\text{–split}}(k_i)$	$\mathrm{rank}_{\mathrm{geom,\,an}}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0.

proof Immediate from Theorem 10.1.7, part 3), and the proof of 8.3.6. QED

Theorem 10.3.4 Suppose we have $\mathcal{F}$ on (C–S)/T satisfying all the hypotheses of Theorem 10.0.9, part 5). Then $\mathcal{G}_\nu$ is orthogonally self–dual toward $\overline{\mathbb{Q}}_\ell(-w-1)$. Fix ν, fix a sign $\varepsilon = \pm 1$, and fix a conjugacy class σ_ν in the symmetric group S_{d_ν}. Take any sequence of data

$$(k_i, t_i, \alpha_{\nu,k_i,t_i})$$

with

k_i a finite field, $\#k_i > \mathrm{Max}(4A(X_\nu/T)^2, 4C(X_\nu/T, \mathcal{S}_{S_{d_\nu}})^2(d_\nu!)^4)$

t_i a k_i–valued point T, with $A(\mathrm{Frob}_{k_i,t_i}) = \varepsilon$,

α_{ν,k_i,t_i} in $(\overline{\mathbb{Q}}_\ell)^\times$ is either choice of $\pm(\#k_i)^{(-w-1)/2}$,

in which $i \mapsto \#k_i$ is strictly increasing. We have the following table of limit formulas. In these tables, the number in the third column is the limit, as $i \to \infty$, of the average value of the quantity in the second column over all f's in the set named in the first column.

ε = +1

$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0,
$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0,
$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	0.

ε = −1

$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{quad,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	1,
$X_{\nu,t_i,\sigma_\nu\text{-split}}(k_i)$	$\mathrm{rank}_{geom,\ an}(\mathcal{G}_{\nu,t_i}, k_i, f)$	2.

proof Immediate from Theorem 10.1.7, part 4), and the proof of 8.3.8. QED

10.4 Interlude: Review of GUE and eigenvalue location measures

(10.4.1) Fix an integer $r \geq 1$ and an offset vector $c = (c(1), ..., c(r))$ in $\mathbb{Z}^r$:

$$0 < c(1) < c(2) < ... < c(r).$$

Define $c(0) := 0$. Given an integer $N > c(r)$, a closed subgroup K of U(N), and an element A in K, write the eigenvalues of A as $e^{i\varphi(j)}$ with angles $\varphi(j)$, j = 1 to N lying in $[0, 2\pi)$:

$$0 \leq \varphi(1) \leq \varphi(2) \leq ... \leq \varphi(N) < 2\pi.$$

Then extend the definition of $\varphi(j)$ to all integers j by requiring

$$\varphi(j + N) = \varphi(j) + 2\pi.$$

From the angles $\varphi(j)$, we next define spacing vectors in $\mathbb{R}^r$. For k = 1 to N, the k'th spacing vector with offsets c attached to A, denoted $s_k(\text{offsets } c)$, is the vector in $\mathbb{R}^r$ whose i'th component is

$$(N/2\pi)(\varphi(k + c(i)) - \varphi(k + c(i-1))).$$

The Borel probability measure on $\mathbb{R}^r$

$$\mu(A, K, \text{offsets } c)$$

is defined to be

$$(1/N)\Sigma_{k = 1 \text{ to } N}\ (\text{delta measure at } s_k(\text{offsets } c)),$$

cf. [Ka–Sa, RMFEM, 1.0].

(10.4.2) For any nonvoid open set K_0 of K, one can make sense of the expected value $\mu(K_0, \text{offsets } c)$ of these measures $\mu(A, K, \text{offsets } c)$ as A varies over K. Formally,

$$\mu(K_0, \text{offsets } c) := \int_{K_0} \mu(A, K, \text{offsets } c)dA,$$

where dA denotes the Haar measure on K, normalized to give K_0 measure one, cf. [Ka–Sa, RMFEM, 1.1]. This expected value measure is a Borel probability measure on $\mathbb{R}^r$.

(10.4.3) The GUE measure μ(univ, offsets c) is the Borel probability measure on $\mathbb{R}^r$ which is the large N limit of the measures μ(U(N), offsets c), cf. [Ka–Sa, RMFEM, 1.2.1] for the precise statement. The universality of μ(univ, offsets c) is this. For each large N separately take $H(N) \subset U(N)$ to be any of

1) any closed subgroup with $SU(N) \subset H(N) \subset U(N)$,
2) any closed subgroup with $SO(N) \subset H(N) \subset U(1)\cdot O(N)$,
3) $O_-(N)$,
4) any closed subgroup with $USp(N) \subset H(N) \subset U(1)\cdot USp(N)$.

Then μ(U(N), offsets c) is the large N limit of the measures μ(H(N), offsets c), cf. [Ka–Sa, RMFEM, 1.2.3 and 1.2.6] for a precise statement.

(10.4.4) The definition of the eigenvalue location measures $\nu(c)$, $\nu(-, c)$, and $\nu(+, c)$ on $\mathbb{R}^r$ attached to the offset vector c is more involved, and requires a case by case discussion.

(10.4.5) To define $\nu(c)$, we begin with U(N) for large N. Given A in U(N), again write its eigenvalues as $e^{i\varphi(j)}$ with angles $\varphi(j) = \varphi(j)(A)$, j = 1 to N lying in $[0, 2\pi)$:

$$0 \le \varphi(1) \le \varphi(2) \le \ldots \le \varphi(N) < 2\pi.$$

Define the normalized angles $\theta(j)(A)$ of A to be the real numbers in $[0, N)$ defined by

$$\theta(j)(A) := (N/2\pi)\varphi(j)(A), \text{ for } j = 1 \text{ to } N.$$

Define a map

$$F_c : U(N) \to \mathbb{R}^r$$

by

$$F_c(A) := (\theta(c(1))(A), \theta(c(2))(A), \ldots, \theta(c(r))(A)).$$

Then we define the Borel probability measure $\nu(U(N), c)$ on $\mathbb{R}^r$ to be the direct image by F_c of the total mass one Haar measure on U(N):

$$\nu(U(N), c) := F_{c*}(\text{total mass one Haar measure on } U(N)).$$

(10.4.6) Similarly, for any of the closed subgroups $U_n(N)$ between SU(N) and U(N), we define

$$\nu(U_n(N), c) := F_{c*}(\text{total mass one Haar measure on } U_n(N)).$$

If we pick, separately for each large N, H(N) to be either U(N) or some $U_n(N)$, then the large N limit of the measures $\nu(H(N), c)$ exists as a Borel probability measure on $\mathbb{R}^r$, cf [Ka–Sa, RMFEM, AD 4.3 and AD 10.2].

(10.4.7) To define $\nu(\pm, c)$ we need to distinguish yet more cases. Suppose first we look at G(2N) which is either USp(2N) or SO(2N). For both these groups, the eigenvalues of any element A occur in N inverse pairs $e^{\pm i\varphi(j)}$ with angles

$$0 \le \varphi(1) \le \varphi(2) \le \ldots \le \varphi(N) \le \pi.$$

We define the normalized angles

$$\theta(j)(A) := (N/\pi)\varphi(j)(A), \text{ for } j = 1 \text{ to } N.$$

For N > c(r), we again define

$$F_c : G(2N) \to \mathbb{R}^r$$

by

$$F_c(A) := (\theta(c(1))(A), \theta(c(2))(A), ..., \theta(c(r))(A)).$$

Then we define the Borel probability measure $\nu(G(2N), c)$ on $\mathbb{R}^r$ as the direct image by F_c of the total mass one Haar measure on G(2N):

$$\nu(G(2N), c) := F_{c*}(\text{total mass one Haar measure on } G(2N)).$$

(10.4.8) For $O_-(2N)$, every element has both ±1 as eigenvalues. The other 2N−2 eigenvalues occur in N−1 inverse pairs $e^{\pm i\varphi(j)}$ with angles

$$0 \le \varphi(1) \le \varphi(2) \le ... \le \varphi(N-1) \le \pi.$$

We define the normalized angles

$$\theta(j)(A) := (N/\pi)\varphi(j)(A), \text{ for } j = 1 \text{ to } N-1.$$

For N−1 > c(r), we define

$$F_c : O_-(2N) \to \mathbb{R}^r$$

by

$$F_c(A) := (\theta(c(1))(A), \theta(c(2))(A), ..., \theta(c(r))(A)).$$

Then we define the Borel probability measure $\nu(O_-(2N), c)$ on $\mathbb{R}^r$ as

$$\nu(O_-(2N), c) := F_{c*}(\text{total mass one Haar measure on } O_-(2N)).$$

(10.4.9) For $O_\pm(2N+1)$, every element A admits the indicated choice of ±1 as an eigenvalue, and the other 2N eigenvalues occur in N inverse pairs $e^{\pm i\varphi(j)}$ with angles

$$0 \le \varphi(1) \le \varphi(2) \le ... \le \varphi(N) \le \pi.$$

We define the normalized angles

$$\theta(j)(A) := ((N + 1/2)/\pi)\varphi(j)(A), \text{ for } j = 1 \text{ to } N.$$

For N > c(r), we define

$$F_c : O_\pm(2N+1) \to \mathbb{R}^r$$

by

$$F_c(A) := (\theta(c(1))(A), \theta(c(2))(A), ..., \theta(c(r))(A)).$$

Then we define the Borel probability measure $\nu(O_\pm(2N+1), c)$ on $\mathbb{R}^r$ as

$$\nu(O_\pm(2N+1), c) := F_{c*}(\text{total mass one Haar measure on } O_\pm(2N+1)).$$

(10.4.10) Having made the relevant definitions, we can now state the large N limit theorems for these measures. The measures

$$\nu(USp(2N), c)$$

on $\mathbb{R}^r$ have a large N limit, denoted $\nu(-, c)$. To state the result for orthogonal groups, we pass to the $O_{\mathrm{sign}\ \varepsilon}$ notation. For each choice of $\varepsilon = \pm 1$, we put

$$O_{\mathrm{sign}\ \varepsilon}(N) := \{A \text{ in } O(N) \text{ with } \det(-A) = \varepsilon\}.$$

The measures

$$\nu(O_{\mathrm{sign}\ -}(N), c)$$

have the same large N limit $\nu(-, c)$ as the measures $\nu(USp(2N), c)$. The measures

$$\nu(O_{\mathrm{sign}\ +}(N), c)$$

have a large N limit, denoted $\nu(+, c)$, on $\mathbb{R}^r$. All three measures

$$\nu(c),\ \nu(-, c),\ \nu(+, c)$$

are Borel probability measures on $\mathbb{R}^r$ which are absolutely continuous with respect to Lebesgue measure, cf. [Ka–Sa, RMFEM, AD 4.3, AD 4.4.1, and AD 10.2].

10.5 Applications to GUE discrepancy

Theorem 10.5.1 Fix an integer $r \geq 1$ and an offset vector $c = (c(1), ..., c(r))$ in $\mathbb{Z}^r$. Fix an integer κ, $1 \leq \kappa \leq r$, and a surjective linear map

$$\pi: \mathbb{R}^r \to \mathbb{R}^\kappa.$$

Denote

$$\mu := \mu(\text{univ, offsets } c).$$

Suppose we are in one of the first three cases (SL, Sp, O, or strongly SO) of Theorem 10.1.7. In the SL case, assume further that for each ν, the group G_{geom} for $\mathcal{G}_{\nu,t}$ on X_t is constant in t. Denote by N_ν the rank of $\mathcal{G}_\nu$, and denote by $K(N_\nu)$ the closed subgroup of $U(N_\nu)$ which is the chosen compact form of the common value of G_{geom} for all the $\mathcal{G}_{\nu,t}$'s.

Pick any sequence (k_i, t_i) of pairs

$$(\text{a finite field } k_i, \text{ a } k_i\text{–valued point } t_i \text{ of } T)$$

in which $i \mapsto \#k_i$ is a strictly increasing sequence. For each ν, the sets $X_{t_i}(k_i)$ are nonempty for large enough i. For each such i we pick an α_{ν,k_i,t_i} as in (the corresponding case of) 10.1.7 and form the measure

$$\mu(k_i, t_i, \alpha_{\nu,k_i,t_i})$$

on $K(N_\nu)^\#$.

For each ν large enough that $N_\nu > c(r)$, form, for each element A in $K(N_\nu)$, the spacing measure on $\mathbb{R}^r$

$$\mu_\nu(A) := \mu(A, K(N_\nu), \text{offsets } c).$$

Then take its direct image

$$\pi_*\mu_\nu(A)$$

to $\mathbb{R}^{\kappa}$. Form the discrepancy [Ka–Sa, RMFEM, 1.0.10]

$$\mathrm{discrep}(\pi_*\mu, \pi_*\mu_\nu(A))$$

between this measure on $\mathbb{R}^{\kappa}$ and the direct image $\pi_*\mu$ of the GUE measure, and view its formation as a continuous $\mathbb{R}$−valued central function

$$A \mapsto \mathrm{Discrep}(A) := \mathrm{discrep}(\pi_*\mu, \pi_*\mu_\nu(A))$$

on $K(N_\nu)$. Consider the integral

$$\int_{K(N_\nu)} \mathrm{Discrep}(A)d\mu(k_i, t_i, \alpha_{\nu,k_i,t_i,\nu})(A)$$
$$:= (1/\#X_{\nu,t_i}(k_i))\Sigma_{x \text{ in } X_{\nu,t_i}(k_i)} \mathrm{Discrep}(\theta(k_i, t_i, x, \alpha_{\nu,k_i,ti})).$$

Then the double limit

$$\lim_{\nu\to\infty}\lim_{i\to\infty} \int_{K(N_\nu)} \mathrm{Discrep}(A)d\mu(k_i, t_i, \alpha_{k_i,t_i,\nu})(A)$$

vanishes. More precisely, given $\varepsilon > 0$, there exists an explicit constant $N(\varepsilon, r, c\,\pi)$ such that if $N_\nu \geq N(\varepsilon, r, c\,\pi)$, we have

$$\lim_{i\to\infty} \int_{K(N_\nu)} \mathrm{Discrep}(A)d\mu(k_i, t_i, \alpha_{k_i,t_i,\nu})(A) \leq (N_\nu)^{\varepsilon - (1/(2r+4))}.$$

If we are in the Sp case or the O case, we can in addition pick a conjugacy class σ_ν in S_{d_ν} for each ν. Then we can consider the sequences of measures

$$\mu(k_i, t_i, \alpha_{k_i,t_i,\nu}, \sigma_\nu\text{−split}) \text{ on } K(N_\nu)^\#.$$

The above results are also valid for this sequence of measures.

In the O case, we can also make a single choice of sign ε, and so we can consider the two sequences of measures

$$\mu(k_i, t_i, \alpha_{k_i,t_i,\nu}, \text{sign } \varepsilon) \text{ on } O_{\text{sign }\varepsilon}(N_\nu, \mathbb{R})^\#$$

$$\mu(k_i, t_i, \alpha_{k_i,t_i,\nu}, \sigma_\nu\text{−split, sign } \varepsilon) \text{ on } O_{\text{sign }\varepsilon}(N_\nu, \mathbb{R})^\#.$$

The above results are also valid for these sequences of measures.

proof This is immediate from Theorem 10.1.7, thanks to [Ka–Sa, RMFEM, 12.1.3]. QED

Theorem 10.5.2 Fix an integer $r \geq 1$ and an offset vector $c = (c(1), ..., c(r))$ in $\mathbb{Z}^r$. Fix an integer κ, $1 \leq \kappa \leq r$, and a surjective linear map

$$\pi: \mathbb{R}^r \to \mathbb{R}^\kappa.$$

Denote

$$\mu := \mu(\text{univ, offsets c}).$$

Suppose we are in the SO/O case of Theorem 10.1.7. Pick a sign $\varepsilon = \pm 1$. Denote by N_ν the rank of $\mathcal{G}_\nu$.

Pick any sequence (k_i, t_i) of pairs

$$(\text{a finite field } k_i, \text{ a } k_i\text{–valued point } t_i \text{ of } T)$$

in which $i \mapsto \#k_i$ is a strictly increasing sequence and in which $A(\mathrm{Frob}_{k_i,t_i}) = \varepsilon$ for every i. For each ν, the sets $X_{t_i}(k_i)$ are nonempty for large enough i. For each such i we pick an α_{ν,k_i,t_i} as in the SO/O case of 10.1.7, and form the measure

$$\mu(k_i, t_i, \alpha_{\nu,k_i,t_i})$$

on $O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R})^\#$.

For each ν large enough that $N_\nu > c(r)$, form, for each element A in $O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R})$, the spacing measure on $\mathbb{R}^r$

$$\mu_\nu(A) := \mu(A, O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R}), \text{ offsets } c).$$

Then take its direct image

$$\pi_*\mu_\nu(A)$$

to $\mathbb{R}^\kappa$. Form the discrepancy [Ka–Sa, RMFEM, 1.0.10]

$$\mathrm{discrep}(\pi_*\mu, \pi_*\mu_\nu(A))$$

between this measure on $\mathbb{R}^\kappa$ and the direct image $\pi_*\mu$ of the GUE measure, and view its formation as a continuous $\mathbb{R}$–valued central (i.e., invariant by $O(N_\nu, \mathbb{R})$ conjugation) function

$$A \mapsto \mathrm{Discrep}(A) := \mathrm{discrep}(\pi_*\mu, \pi_*\mu_\nu(A))$$

on $O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R})$. Consider the integral

$$\int_{O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R})} \mathrm{Discrep}(A) d\mu(k_i, t_i, \alpha_{\nu,k_i,t_i,\nu})(A)$$
$$:= (1/\#X_{\nu,t_i}(k_i))\Sigma_{x \text{ in } X_{\nu,t_i}(k_i)} \mathrm{Discrep}(\theta(k_i, t_i, x, \alpha_{\nu,k_i,ti})).$$

Then the double limit

$$\lim_{\nu\to\infty}\lim_{i\to\infty} \int_{O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R})} \mathrm{Discrep}(A) d\mu(k_i, t_i, \alpha_{k_i,t_i,\nu})(A)$$

vanishes. More precisely, given $\varepsilon > 0$, there exists an explicit constant $N(\varepsilon, r, c\ \pi)$ such that if $N_\nu \ge N(\varepsilon, r, c\ \pi)$, we have

$$\lim_{i\to\infty} \int_{O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R})} \mathrm{Discrep}(A) d\mu(k_i, t_i, \alpha_{k_i,t_i,\nu})(A) \le (N_\nu)^{\varepsilon - (1/(2r+4))}.$$

Pick a conjugacy class σ_ν in S_{d_ν} for each ν, and consider the sequences of measures

$$\mu(k_i, t_i, \alpha_{k_i,t_i,\nu}, \sigma_\nu\text{–split}) \text{ on } O_{\mathrm{sign}\ \varepsilon}(N_\nu, \mathbb{R})^\#.$$

Then the above results are also valid for this sequence of measures.

proof This is immediate from Theorem 10.1.7, thanks to [Ka–Sa, RMFEM, 12.1.3]. QED

10.6 Application to eigenvalue location measures

Theorem 10.6.1 Fix an integer $r \geq 1$ and an offset vector $c = (c(1), ..., c(r))$ in $\mathbb{Z}^r$. Suppose we are in one of the cases of Theorem 10.1.7. In the SL case, assume further that for each ν, the group G_{geom} for $\mathcal{G}_{\nu,t}$ on X_t is constant in t. In the SO/O case, pick a sign $\varepsilon = \pm 1$.

Pick any sequence (k_i, t_i) of pairs

(a finite field k_i, a k_i–valued point t_i of T)

in which $i \mapsto \#k_i$ is a strictly increasing sequence. If we are in the SO/O case, assume in addition that

$$A(\mathrm{Frob}_{k_i,t_i}) = \varepsilon, \text{ for every i.}$$

For each ν, the sets $X_{t_i}(k_i)$ are nonempty for large enough i, and for each such i we pick an α_{ν,k_i,t_i} as in (the corresponding case of) 10.1.7, and form the measure

$$\mu(k_i, t_i, \alpha_{\nu,k_i,t_i})$$

on

$U_{m_\nu}(N_\nu)^\#$, in the SL case,

$USp(N_\nu)^\#$, in the Sp case,

$O(N_\nu, \mathbb{R})^\#$, in the O case,

$SO(N_\nu, \mathbb{R})^\#$, in the strongly SO case,

$O_{\text{sign } \varepsilon}(N_\nu, \mathbb{R})^\#$, in the SO/O case.

If we are in the Sp case, we can in addition pick a conjugacy class σ_ν in S_{d_ν} for each ν. Then we can consider the sequences of measures

$$\mu(k_i, t_i, \alpha_{k_i,t_i,\nu}, \sigma_\nu\text{–split}) \text{ on } USp(N_\nu)^\#.$$

In the O case, we can pick a conjugacy class σ_ν in S_{d_ν} for each ν, and we can also make a single choice of sign ε. So we can consider the two sequences of measures

$$\mu(k_i, t_i, \alpha_{k_i,t_i,\nu}, \text{sign } \varepsilon) \text{ on } O_{\text{sign } \varepsilon}(N_\nu, \mathbb{R})^\#$$

$$\mu(k_i, t_i, \alpha_{k_i,t_i,\nu}, \sigma_\nu\text{–split, sign } \varepsilon) \text{ on } O_{\text{sign } \varepsilon}(N_\nu, \mathbb{R})^\#.$$

Then we have the following integration formulas. Fix a continuous function h of compact support on $\mathbb{R}^r$.

1) If we are in the SL case, we can compute $\int_{\mathbb{R}^r} h d\nu(c)$ as the double limit

$$\lim_{\nu \to \infty} \lim_{i \to \infty} \int_{U_{m_\nu}(N_\nu)} h(F_c(A)) d\mu(k_i, t_i, \alpha_{\nu,k_i,t_i})(A).$$

2) If we are in the Sp case, we can compute $\int_{\mathbb{R}^r} h d\nu(-,c)$ as the double limit

$$\lim_{\nu \to \infty} \lim_{i \to \infty} \int_{USp(N_\nu)} h(F_c(A)) d\mu(k_i, t_i, \alpha_{\nu, k_i, t_i})(A),$$

or as the double limit

$$\lim_{\nu \to \infty} \lim_{i \to \infty} \int_{USp(N_\nu)} h(F_c(A)) d\mu(k_i, t_i, \alpha_{\nu, k_i, t_i}, \sigma_\nu\text{–split})(A).$$

3) If we are in the O case, then for either choice of sign ε, we can compute $\int_{\mathbb{R}^r} hd\nu(\varepsilon, c)$ as the double limit

$$\lim_{\nu \to \infty} \lim_{i \to \infty} \int_{O_{\text{sign } \varepsilon}(N_\nu)} h(F_c(A)) d\mu(k_i, t_i, \alpha_{\nu, k_i, t_i}, \text{sign } \varepsilon)(A),$$

or as the double limit

$$\lim_{\nu \to \infty} \lim_{i \to \infty} \int_{O_{\text{sign } \varepsilon}(N_\nu)} h(F_c(A)) d\mu(k_i, t_i, \alpha_{\nu, k_i, t_i}, \sigma_\nu\text{–split}, \text{sign } \varepsilon)(A).$$

4) If we are in the SO/O case, and have chosen the sign ε, we can compute $\int_{\mathbb{R}^r} hd\nu(\varepsilon, c)$ as the double limit

$$\lim_{\nu \to \infty} \lim_{i \to \infty} \int_{O_{\text{sign } \varepsilon}(N_\nu)} h(F_c(A)) d\mu(k_i, t_i, \alpha_{\nu, k_i, t_i})(A),$$

or as the double limit

$$\lim_{\nu \to \infty} \lim_{i \to \infty} \int_{O_{\text{sign } \varepsilon}(N_\nu)} h(F_c(A)) d\mu(k_i, t_i, \alpha_{\nu, k_i, t_i}, \sigma_\nu\text{–split})(A).$$

proof This is immediate from Theorem 10.1.7, thanks to [Ka–Sa, RMFEM, AD 4.3, AD 10.2, and AD 11.4]. QED

References

[Beu–St] Beukers, F., and Stienstra, J., On the Picard–Fuchs equations and the formal Brauer group of certain elliptic K3–surfaces, Math. Ann. 271 (1985), no. 2, 269–304.

[Bir–St] Birch, B. J., and Stephens, N. M., The parity of the rank of the Mordell–Weil group, Topology 5 (1966), 295–299.

[Blich–FCG] Blichfeldt, H. F., *Finite Collineation Groups*, U. Chicago Press, 1917.

[Br–Con–Dia–Tay] Breuil, C., Conrad, B., Diamond, F., and Taylor, R., On the modularity of elliptic curves over $\mathbb{Q}$: wild 3–adic exercises, J. Amer. Math. Soc. 14 (2001), no. 4, 843–939.

[Brum–McG] Brumer, A., and McGuinness, O., The behavior of the Mordell–Weil group of elliptic curves, BAMS (N.S.) 23 (1990), 375–382.

[Brum] Brumer, A., The average rank of elliptic curves I, Invent. Math. 109 (1992), 445–472.

[deJ–Ka] de Jong, A. J., and Katz, N., Monodromy and the Tate conjecture: Picard numbers and Mordell–Weil ranks in families, Israel J. Math. 120 (2000), part A, 47–79.

[De–AFT] Deligne, P., Application de la formule des traces aux sommes trigonométriques, in *Cohomologie Etale (SGA 4 1/2)*, Springer Lecture Notes in Mathematics 569, 1977, 168–232.

[De–Const] Deligne, P., Les constantes des equations fonctionelles des fonctions L, in *Modular Functions of One Variable III*, Springer Lecture Notes in Mathematics 349, 1973, 501–597.

[De–Weil II] Deligne, P., La conjecture de Weil II, Pub. Math. I.H.E.S. 52 (1981), 313–428.

[EGA] Grothendieck, A., rédigé avec la collaboration de Dieudonné, J., *Elements de Géométrie Algébrique*, Pub. Math. I.H.E.S. 4, 8, 11, 17, 20, 24, 28, 32, 1960–1967.

[Fer–22] Fermigier, S., Une courbe elliptique définie sur $\mathbb{Q}$ de rang ≥ 22, Acta Arith. 82 (1997), 359–363.

[Fer–EE] Fermigier, S., Étude expérimentale du rang de famillies de courbes elliptiques sur $\mathbb{Q}$, Exper. Math. 5 (1966), 119–130.

[Go] Goldfeld, D., Conjectures on elliptic curves over quadratic fields, in *Number Theory, Carbondale 1979*, Springer Lecture Notes in Mathematics 751, 1979, 108–118.

[Gou–Maz] Gouvea F., and Mazur, B., The square–free sieve and the rank of elliptic curves, JAMS 4 (1991), 1–23.

[Harb–AC] Harbater, D., Abhyankar's conjecture on Galois groups over curves, Invent. Math. 117 (1994), 1–25.

[Harb–Mod] Harbater, D., Moduli of p–covers of curves, Comm. Algebra 8 (1980), 1095–1122.

[Hart] Hartshorne, R., *Algebraic Geometry*, Springer Graduate Texts in Mathematics 52, 1977.

[Hasse] Hasse, H., Beweis der Analogons der Riemannschen Vermutung für die Artinschen und F.K. Schmidschen Kongruenze–zetafunktionen in gewissen eliptischen Fällen, Ges. d. Wiss. Nachrichten Math–Phys. Klasse, Heft 3 (1933), 253–262.

[Huf–Wa] Huffman, W. C., and Wales, D. B., Linear groups of degree n containing an element with exactly n–2 equal eigenvalues, Linear and Multilinear Algebra 5 (1975), 53–59.

[Ka–ACT] Katz, N., Affine cohomological transforms, perversity and monodromy, JAMS 6 (1993), 149–222.

[Ka–ASDE] Katz, N., Algebraic solutions of differential equations; p–curvature and the Hodge filtration, Inv. Math. 18 (1972), 1–118.

[Ka–ESDE] Katz, N., *Exponential Sums and Differential Equations*, Annals of Math. Study 124, Princeton Univ. Press, 1990.

[Ka–GKM] Katz, N., *Gauss Sums, Kloosterman Sums, and Monodromy Groups,* Annals of Math. Study 116, Princeton Univ. Press, 1988.

[Ka–LG] Katz, N., Local to global extensions of representations of fundamental groups, Ann. Inst. Fourier 36, 4 (1986), 59–106.

[Ka–MG] Katz, N., On the monodromy groups attached to certain families of exponential sums, Duke Math. J. 54 No. 1, (1987), 41–56.

[Ka–NCMT] Katz, N., Nilpotent connections and the monodromy theorem: applications of a result of Turritin, Pub. Math. I.H.E.S. 39 (1971), 175–232.

[Ka–RLS] Katz, N., *Rigid Local Systems*, Annals of Math. Study 138, Princeton Univ. Press, 1995.

[Ka–SE] Katz, N., *Sommes Exponentielles*, rédigé par G. Laumon, Asterisque 79, 1980.

[Ka–Spacefill] Katz, N., Spacefilling curves over finite fields, Math. Res. Letters 6 (1999), 613–624.

[Ka–Maz] Katz, N., and Mazur, B., *Arithmetic Moduli of Elliptic Curves*, Annals of Math. Study 108, Princeton Univ. Press, 1985.

[Ka–Sar, RMFEM] Katz, N., and Sarnak, P., *Random Matrices, Frobenius Eigenvalues, and Monodromy*, A.M.S. Colloquium Publications 45, 1999.

[Ka–Sar, Zeroes] Katz, N., and Sarnak, P., Zeroes of zeta functions and symmetry, BAMS 36 (1999), 1–26.

[Kra–Zag] Kramarz, G., and Zagier, D., Numerical investigations related to the L–series of certain elliptic curves, J. Indian Math. Soc. 97 (1987), 313–324.

[Lang] Lang, S., *Algebra, Third Edition*, Addison–Wesley, 1993.

[Lau–SC] Laumon, G., Semi–continuité du conducteur de Swan (d'après P. Deligne), in *Caractéristique d'Euler–Poincaré, Seminaire E.N.S. 1978–79*, Astérisque 82–83 (1981), 173–219.

[Lau–TFC] Laumon, G., Transformation de Fourier, constantes d'équations fonctionelles et conjecture de Weil, Pub. Math. I.H.E.S. 65 (1987), 131–210.

[Mar–McM–23] Martin, R., and McMillen, W., An Elliptic Curve over Q with Rank at least 23, preprint, June 1997, posted to sci.math.numberthy by Victor Miller on 16 Mar 98 16:14:56 GMT.

[Mar–McM–24] Martin, R., and McMillen, W., An Elliptic Curve over Q with Rank at least 24, preprint, January 2000, posted to NmbrThry@LISTSERV.NODAK.EDU by Victor Miller on 2 May 2000 13:30:42.

[Mes–11] Mestre, J.–F., Courbes elliptiques de rang ≥ 11 sur $\mathbb{Q}(t)$, C. R. Acad. Sci. Paris Sér. I Math. 313 (1991), no. 3, 139–142.

[Mes–12] Mestre, J.–F., Courbes elliptiques de rang ≥ 12 sur $\mathbb{Q}(t)$, C. R. Acad. Sci. Paris Sér. I Math. 313 (1991), no. 4, 171–174.

[Mes–15] Mestre, J.–F., Un exemple de courbe elliptique sur $\mathbb{Q}$ de rang ≥ 15, C. R. Acad. Sci. Paris Sér. I Math. 314 (1992), no. 6, 453–455.

[Mit] Mitchell, H., Determination of all primitive collineation groups in more than four variables which contain homologies, Amer. J. Math. 36 (1914), 1–12.

[Mor] Mordell, L. J., On the rational solutions of the indeterminate equations of the 3rd and 4th degrees, Proc. Camb. Phil. Soc. 21 (1922), 179–192.

[Mum–TCAC] Mumford, D., Theta characteristics of an algebraic curve. Ann. Sci. École Norm. Sup. (4) 4 (1971), 181–192.

[Nag–20] Nagao, K., An example of elliptic curve over Q with rank 20, Proc. Japan Acad., Ser. A, 69, No. 8 (1993), 291–293.

[Nag–Ko–21] Nagao, K., and Kouya, T., An example of elliptic curve over Q with rank 21, Proc. Japan Acad. Ser. A, 70, No. 4 (1994), 104–105.

[Ner–10] Néron, A., Un théorème sur le rang des courbes algébriques dans les corps de degré de transcendance fini, C. R. Acad. Sci. Paris 228 (1949), 1087–1089.

[SGA] Grothendieck, A., et al., *Séminaire de Géométrie Algébrique du Bois–Marie*, SGA 1, SGA 4 Parts I, II, and III, SGA 4 1/2, SGA 5, SGA 7 Parts I and II, Springer Lecture Notes in Math. 224, 269–270–305, 569, 589, 288–340, 1971 to 1977.

[S–T, GR] Serre, J.–P., and Tate, J., Good reduction of abelian varieties, Ann. Math. (2) 88 (1968), 492–517.

[Shio] Shioda, T., Some remarks on elliptic curves over function fields, in *Journées Arithmétiques, 1991 (Geneva)*, Astérisque 209, 1992, 99–114.

[Shio–In] Shioda, T., and Inose, H., On singular K3 surfaces, in *Complex Analysis and Algebraic Geometry*, Iwanami Shoten, Tokyo, 1977, 119–136.

[Sil–ATEC] Silverman, J. H., *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer Graduate Texts in Mathematics 151, 1994.

[Ta–Alg] Tate, J., Algorithm for determining the type of a singular fiber in an elliptic pencil, in *Modular Functions of One Variable IV (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, Springer Lecture Notes in Mathematics 476, 1975, 33–52.

[Ta–BSD] Tate, J., On the conjectures of Birch and Swinnerton–Dyer and a geometric analog, in *Séminaire Bourbaki*, Vol. 9, Exp. No. 306, 415–440, Soc. Math. France, Paris, 1995.

[Tay–Wi] Taylor, R., and Wiles, A., Ring–theoretic properties of certain Hecke algebras, Ann. Math. (2) 141 (1995), 553–572.

[Th–QP] Thompson, J. G., Quadratic pairs, in *Actes du Congres International des Mathématiciens (Nice, 1970)*, Tome 1, Gauthier–Villars, Paris, 1971, 375–376.

[Wa–Ta] Wada, H., and Taira, M., Computations of the rank of elliptic curves $y^2 = x^3 - n^2x$, Proc. Japan Acad. Ser. A Math. Sci. 70 (1994), 154–157.

[Wales] Wales, D., Quasiprimitive linear groups with quadratic elements, J. Alg., to appear.

[Wi] Wiles, A., Modular elliptic curves and Fermat's last theorem, Ann. Math. (2) 141 (1995), 443–551.

[Zal] Zalesskii , A. E., Linear groups, Russian Math Surveys 36:5 (1981), 63–128.

[Zar–SLA] Zarhin, Yu. G., Linear simple Lie algebras and ranks of operators, in *The Grothendieck Festschrift, Vol. III*, Progr. Math. 88, Birkhäuser, 1990, 481–495.

Index

[The index lists page numbers preceded by section numbers, except for page numbers 3–21 (the pages of the Introduction). Thus "Abhyankar Conjecture" is found on page 155, in section (8.4.3), while "average rank" occurs on pages 3–4, 6, 8, 10, 15–17, 222 (in section (10.3)), and 224 (in section (10.3.2)). When several pages are cited for an item, the most important, if any, are in **boldface** .]